The spatial scale of adaptation in common ragweed (*Ambrosia artemisiifolia*)

A Dissertation
SUBMITTED TO THE FACULTY OF
THE UNIVERSITY OF MINNESOTA
BY

Amanda Jane Gorton

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Advisors: David A. Moeller, Peter Tiffin

September 2019

Copyright Page

Acknowledgments

First and foremost, thank you to my advisors, Peter Tiffin and David Moeller, for their thoughtful insights and guidance on my dissertation research. I am deeply indebted to them for their contributions to multiple aspects of my dissertation including problem solving, experimental design, and writing. I grateful for their willingness to let me pursue my interests and advise me on a plant system that was not part of either of their research programs.

I would also like to thank my committee members, Ruth Shaw and Rebecca Montgomery, for years of helpful feedback. In addition, I am grateful for the expertise and contributions of Tuomas Hämälä on Chapter 3, and advice from Jennifer Powers and Leland Werden on rainout shelter design in Chapter 2. Thank you to Grinnell College, the University of Western Illinois, Washington University, and the University of Minnesota for providing the space for the field sites in Chapter 4. None of my dissertation research would have been possible without the generous funding I received from the National Science and Engineering Research Council of Canada, the Graduate School at University of Minnesota, the University of Minnesota College of Biological Sciences, the Bell Natural History Museum, the Society for the Study of Evolution, and the Department of Ecology, Evolution and Behavior.

I am very fortunate to have had the help and support of colleagues, friends and family throughout the six years of my dissertation. In particular, I would like to thank Mary Merello (Missouri Botanical Garden), who provided assistance and companionship with much of my seed collections. Ryan Briscoe Runquist, Liana Burghardt, and Mason Kulbaba all provided invaluable help with my data analyses, and were additional sources of mentorship. I was also lucky to have the help with data collection in the field from several hardworking undergraduates at the University of Minnesota, in particular Eric Holton and Joo Yoon Kim. To my fantastic, supportive, and brilliant friends: I am forever grateful for your help, emotional support, and advice throughout my graduate degree. In particular, thank you Emmie Miller, Marta Lyons, John Benning, Anna Peschel, Susy Ireland, Leigha Dimitroff, Anika Bratt, and Josie Griffin. A special thank you to Ann

Willis, who drove up to Itasca and took care of Hattie for an entire week so that I could write. Finally, to my husband, life partner, and fellow scientist, Charlie Willis: thank you for your numerous intellectual contributions to my dissertation over the years, taking care of our new family while I finished, and for your love, support, and unwavering belief in me.

Dedication

I dedicate my dissertation to my daughter, Harriet Ann Gorton Willis, who provided me with both the motivation to finish my dissertation and much needed distraction after a long day of writing.

Table of Contents

List of Tables

List of Figures

Chapter 1

Little plant big city: a test of adaptation to urban environments in common ragweed (*Ambrosia artemisiifolia*)

ABSTRACT

A full understanding of how cities shape adaptation requires characterizing genetically-based phenotypic and fitness differences between urban and rural populations under field conditions. We used a reciprocal transplant experiment with the native plant common ragweed, (*Ambrosia artemisiifolia*), and found that urban and rural populations have diverged in flowering time, a trait that strongly affects fitness. Although urban populations flowered earlier than rural populations, plants growing in urban field sites flowered later than plants in rural field sites. This counter-gradient variation is consistent adaptive divergence between urban and rural populations. Also consistent with local adaptation, both urban and rural genotypes experienced stronger net selection in the foreign than in the local habitat, but this pattern was not significant for male fitness. Despite the evidence for local adaptation, rural populations had higher lifetime fitness at all sites, suggesting that selection has been stronger or more uniform in rural than urban populations. We also found that inter-population differences in both flowering time and fitness tended to be greater among urban than rural populations, which is consistent with greater drift or spatial variation in selection within urban environments. In sum, our results are consistent with adaptive divergence of urban and rural populations, but also suggest there may be greater environmental heterogeneity in urban environments which also affects evolution in urban landscapes.

Published as:

Gorton, A. J., D. A. Moeller, and P. Tiffin. 2018. Little plant, big city: a test of adaptation to urban environments in common ragweed (*Ambrosia artemisiifolia*). Proceedings of the Royal Society B Biological Sciences 285:20180968.

INTRODUCTION

Urbanization causes pronounced changes in the biotic and abiotic environment. Urban areas are warmer (Oke 1973), have increased levels of CO_2 and ozone (Fenger 1999), greater salinity stress (Cunningham et al. 2007) and altered precipitation patterns (Huff and Changnon 1973) compared to non-urban areas. These environmental changes can reduce biodiversity (McKinney 2008; Aronson et al. 2014), modify plant growth and physiology (Gregg et al. 2003), and change animal behaviors (Slabbekoorn and Peet 2003). The evolutionary consequences of urbanization have been studied far less than these ecological consequences (Alberti 2015; Johnson and Munshi-South 2017). Nevertheless, urban environments are likely to alter both non-adaptive and adaptive evolutionary processes (Johnson and Munshi-South 2017). Urban populations are often smaller and more fragmented than rural populations, thereby increasing genetic drift and genetic differentiation among populations (Munshi-South and Kharchenko 2010; Lourenço et al. 2017). Moreover, environmental differences within and between urban and rural environments may cause spatially-varying natural selection and adaptive divergence among populations.

The results of a handful of studies on animals and plants are suggestive of adaptive divergence between urban and rural populations (Nacci et al. 2010; Atwell et al. 2012; Reid et al. 2016; Winchell et al. 2016). Three studies have investigated urban-rural divergence of plant populations (Cheptou et al. 2008; Thompson et al. 2016; Yakub and Tiffin 2016). Those studies found evolutionary changes in dispersal (Cheptou et al. 2008), flowering time and fecundity (Yakub and Tiffin 2016) and chemical defense (Thompson et al. 2016). All three studies indicate that urban plant populations have phenotypically diverged from rural populations. However, only one of these (Thompson et al. 2016), was conducted under field conditions, limiting our ability to characterize how selection is acting in contemporary populations.

Studies investigating adaptation to urban environments have generally treated urban areas as a single environment, effectively ignoring the diversity of habitats found within cities (*e.g.,* parks, roadsides, mowing, sidewalks, railroad tracks). This environmental heterogeneity, which has been discussed extensively in the urban ecology

literature (Cadenasso et al. 2007; Pickett et al. 2016), may lead to differing selection pressures and adaptive divergence among subpopulations found within a single urban area. Although several studies have investigated molecular divergence among urban populations (Munshi-South and Kharchenko 2010; Serieys et al. 2015; Lourenço et al. 2017), the extent to which subpopulations from a single urban area have phenotypically diverged from one another remains unexplored.

Our primary objective of this study was to advance our understanding of local adaptation to urban environments in plants using realistic field experiments. Our specific objectives were to: i) characterize whether urban and rural populations differ in fitness and phenotypic traits that are likely subject to strong selection; ii) compare the extent of inter-population phenotypic divergence among urban and among rural populations. iii) determine whether selection acting on these traits differs between urban and rural environments or between urban and rural genotypes; and iv) test for evidence of local adaptation to urban and rural environments. If urban and rural populations are both locally adapted to their home environments, we expect phenotypic divergence at the selected traits, different patterns of selection between urban and rural environments, weaker selection acting on local genotypes compared to foreign genotypes, and that both urban and rural populations will have higher fitness when grown in their home environment. We compare the extent of inter-population phenotypic divergence among urban and rural populations to gain insight to how environmental heterogeneity within these broadly defined environments may affect evolution.

To achieve these objectives, we used a multi-site (two rural and two urban sites) reciprocal transplant experiment with the plant *Ambrosia artemisiifolia* (common ragweed, Asteraceae) to conduct the most comprehensive study of local adaptation to urban environments in plants to date. At each site, we collected data on whole-organism phenotypes and lifetime fitness and used these data to characterize selection acting on local and foreign genotypes. We chose *A. artemisiifolia* because it is native to North America, is widely abundant in urban areas, and is an annual, which makes it feasible to conduct experimental manipulations and collect data on lifetime fitness.

METHODS

Study System

We collected seeds from multiple urban and rural populations of common ragweed (*Ambrosia artemisiifolia* L.) in the Minneapolis – Saint Paul, Minnesota, USA metropolitan area. We planted these field collected seeds in four common gardens and visited each site regularly to collect data on phenological traits and fitness. *A. artemisiifolia* is an annual, self-incompatible (Friedman and Barrett 2008), monoecious, and wind-pollinated (Jones 1936; Essl et al. 2015), early-successional species that is native to and widely distributed in North America, and is invasive in Europe (Chauvel et al. 2006), Asia (Xu et al. 2006), and Australia (Bass et al. 2000). The species is abundant in urban areas and in marginal and disturbed habitats (Bassett and Crompton 1975; Marks 1983).

Staminate capitula (i.e. male flowering heads) are in spike-like racemes, hereafter referred to as 'male flowers', which produce pollen that is one of the primary causes of summer and fall allergic rhinitis (Lewis et al. 1983). Pistillate capitula (i.e. female flowering heads) are found in axillary clusters below the male flowers. Individual flowers develop into achenes (small, single seeded fruit) which readily falls off the plant once ripe. The groups of achenes form from all flowers on each capitula are hereafter referred to as 'fruits'.

The Minneapolis – Saint Paul metropolitan area has a population of over 3.5 million people with a high-density urban core. Adjacent rural environments are primarily agricultural, with some forest, prairie and wetlands. From 2011-2014 the urban core was on average 2.5°C warmer than surrounding rural areas (Smoliak et al. 2015).

Seed collections

In 2014, we collected seeds from eight urban and eight rural populations from within 55 km of downtown (rural area classified as agricultural or pastoral (Homer et al. 2015), Figure 1; Appendix 1: Table S1, S2). Seeds were collected from individual maternal plants, each separated by at least 3 meters, at each of the 16 collection locales. Seeds from each maternal plant were kept separate from seeds from other maternal

plants. Across both regions (ca. 50 km^2), we stratified our sampling to broadly sample environments; the distance between adjacent populations ranged from 1 to 8 km. The urban populations were collected in the downtown core of Minneapolis which has a high concentration of buildings and impervious surfaces (Appendix 1: Table S2). The urban populations were collected from a range of habitats including bike paths, road medians, abandoned lots and riverbanks. The rural populations were collected from the edges of agricultural fields and in roadside ditches that were outside of the urban heat island (Smoliak et al. 2015).

Reciprocal transplant experiment

In Spring 2015, we planted four common gardens: two in urban areas and two in rural areas (Figure 1). In each region (*i.e.,* urban, rural), we established one large (urban: USC; rural: RRO) and one small common garden (urban: USP rural: RRT). Hereafter, the common gardens are referred to as 'sites'. Except for RRT, which has a sandy loam soil, the field sites have soil that consist of >60% clay. All sites were sprayed with herbicide (Roundup, Monsanto, MO) on 19 May, 2015 to remove pre-exiting vegetation and we tilled the soil on 22 May, 2015 to create an open, highly disturbed environment, which mimics the natural growing conditions of *A. artemisiifolia*.

At each site, we transplanted approximately four seedlings from each of 8-14 maternal families per population (Appendix 1: Table S1; average number of seeds germinated per line = 7.5 $\pm$ 0.5 SE). All seeds from each maternal family (*i.e.,* genotype) were weighed to obtain an average weight per seed per maternal family. Seeds were stratified in mesh bags by burying them in moist silica sand and keeping them in the dark at 4°C for 10 weeks. After stratification, seeds were planted in 72-cell trays in a 50:50 mix of local field soil and Sunshine Mix #1 (Sun Gro, MA), and germinated in the greenhouse under a 14-hour day and at a day/night temperature of 22/20°C. After two weeks, we transplanted seedlings directly into the soil at each field site (from 28 May to 3 June, USC: n = 528; USP: n = 228; RRO: n = 532; RRT: n = 256). This matches the phenology and approximate size of seedlings previously observed in urban Minneapolis (Gorton, personal observation). For each site, we arranged the seedlings in a completely

randomized design with plants spaced 15 cm apart along each of 10 rows, with the distance between columns alternating between 30 and 80 cm. Seedlings were watered immediately after transplanting and once during the first week after transplanting. The plants were not fertilized during the experiment.

Data collection

We visited each site twice weekly to collect data on three phenological traits: date of transition to reproductive phase, date of first male flower and date of first female flower. The transition to reproductive phase was scored as the date on which a reproductive bud at the apical meristem first appeared. The first male flower was scored as the date on which the first anther opened and shed pollen (*i.e.* the first open male flower). The first female flower was scored as the date at which stigmas first appeared. We also calculated the time between first open male flower and first open female flower (male to female flower). As a proxy for final size, we measured height for each individual between 21 September and 5 October. This is after individuals stopped growing and the weeks during which there is a 50% chance of a frost, based on data from 1981-2010 (0°C, U.S. Climate Normals, NOAA). On the same date, we estimated fitness by conducting flower and fruit counts on a subsample of branches on each plant. Gorton counted the number of male flowers and number of fruits on every fourth branch of each plant (ca. one fourth of each plant), starting with the lowest and largest branch of each plant, and moving up towards the apical meristem. Each plant had a minimum of 20 branches. We multiplied this number by four to get a whole-plant estimate of male and female fitness. We recognize that this is an approximation of fitness and subsampling branches may have introduced variance in our estimates that weaken the statistical power. Nevertheless, due to the size of the plants more precise measures of fitness, e.g. counts every flower and fruit on every branch of every plant, were not feasible.

Statistical analyses of phenotypic divergence and adaptation

We tested for phenotypic differences at two spatial scales. At the regional level, we tested for phenotypic differences between urban and rural seed sources, ignoring

source population. At the population level, we tested for phenotypic differences among source populations within either the urban or rural region. These tests of regional and population differentiation were conducted using data from each of the four field sites. All model equations for the statistical analyses conducted below are included in the supplementary text A and all data analyses were conducted in R, version 3.2.2 (R Core Team 2013).

To test for phenology, size, and fitness differences among sites or between urban and rural seed sources (hereafter 'source region'), we fit linear mixed model (LMMs, *lme4* package (Bates et al. 2015)). We ran separate models with each of four phenology traits (date of transition to reproductive phase, date of first male flower, date of first female flowers, male to female flower) as the dependent variables, average seed weight as a covariate, site, and source region as fixed effects, and the interaction between site and source region, and maternal plant (the plant from which seed was collected) as a random effect (Appendix 1: Supplementary text, A1). Average seed weight was included as a covariate to account for potential maternal effects, as we used field collected seeds.

For the fitness and size analyses, we scaled male and female fitness and height within each site using z-scores (trait value – site mean)/site standard deviation. The z-score transformation was used because plants at the USP field site were considerably larger than plants at the other field sites (mean height in cm: USP = 129.1, USC = 61.5, RRO = 79.2, RRT = 55.6, LRT: $p < 0.0001$ for site), and significant interactions can be caused by both change in mean and change in variance (Yamada 1962).

When these initial analyses for fitness and size revealed a significant or marginally significant site × source region interaction, we conducted separate analyses on the urban and rural seed sources within each site. For each site, we used generalized linear mixed models (GLMMs, *lme4* package, (Bates et al. 2015)) with a Poisson error distribution and a log link function for the fitness variables, and LMMs as above for the size traits. Average seed weight and source region were modeled as fixed effects, and maternal plant as a random effect (Appendix 1: Supplementary text, A2).

For all analyses, we determined the significance of fixed effects and interactions by comparing sequential nested models with and without the term of interest using

likelihood ratio (LR) tests. To simplify analyses, interaction terms with $p > 0.1$ were dropped from the models, but all main effects were retained. Least-square means were extracted from each model using the *lsmeans* package (Lenth 2016). For both male and female fitness, we also calculated local-foreign contrasts (Kawecki and Ebert 2004) within each site using these least-square means (contrast = local source mean – foreign source mean).

To test for phenology and fitness differences among urban populations and among rural populations, we conducted two series of analyses: one analysis was conducted using data from only urban populations and one using data from only rural populations. These analyses were similar to the urban vs rural analyses, but instead of source region in the model, we used 'source population' (Appendix 1: Supplementary text, A3) and maternal plant was nested in population as a random effect. In cases where a significant effect of population was found, we used Tukey tests implemented in the *multcomp* package (Hothorn et al. 2008) to conduct pairwise comparisons among populations. In addition, we conducted Levene's test (*levene.test*, *car* package (Fox and Weisberg 2011)) on the least-square means to determine whether the variance among urban populations and among rural populations was equal.

Aster analysis

We also used 'aster' (*aster* package (Geyer et al. 2007; Shaw et al. 2008)) to compare lifetime fitness of urban and rural populations. Aster explicitly models the dependence of fitness components expressed later in development (*e.g.* fecundity) on those expressed earlier (*e.g.* survival) and allow for different statistical distributions for each fitness component (Geyer et al. 2007; Shaw et al. 2008). As such, aster models can be more statistically powerful than traditional GLMMs. For each individual, we used the following graphical model to estimate lifetime female fitness (number of fruits):

$$\mathbf{1} \rightarrow \underset{(0,1)\ Bernoulli}{\boldsymbol{Surivval\ to\ 6\ weeks}} \rightarrow \underset{(0,1)\ Bernoulli}{\boldsymbol{Flowering}} \rightarrow \underset{Poisson}{\boldsymbol{Number\ of\ fruit}}$$

The same model was used for male fitness, except that number of male flowers replaced number of fruits. In addition to the graphical model, our aster analyses included source region, site, and source region × site as fixed effects. We determined the significance of

fixed effects and interactions by comparing sequentially nested models with and without the term of interest using LR tests. We predicted male and female lifetime fitness separately for each source region at each site, and calculated local-foreign contrasts (contrast = local source mean – foreign source mean) with these predicted means.

Selection analyses

We tested whether selection differs between urban and rural genotypes at each field site and whether foreign genotypes experienced stronger selection than local genotypes, a pattern consistent with local adaptation. We conducted the selection analyses on four traits (transition to reproductive phase, date of first open male flower, time between first male and first female flower, and height), some of which were highly correlated (Appendix 1: Table S11, S12). These traits are all key contributors to plant fitness and are often under strong selection (Hall and Willis 2006; Franks et al. 2007). We estimated selection differentials (net selection) and selection gradients (direct selection) using simple and multiple linear regression, respectively, of male and female relative fitness on standardized traits (Lande and Arnold 1983). Selection differentials measure selection via both direct selection acting on the trait of interest and selection acting through correlated traits; whereas selection gradients provide estimates of selection on a trait after statistically removing indirect selection that results from selection acting on other measured, but not unmeasured, traits. We calculated relative fitness by dividing individual male and female fitness by the site mean. We standardized traits within each site by subtracting the site mean and dividing by the standard deviation. Last, we calculated the genotypic means for relativized fitness metrics and standardized traits. The significance of differentials and gradients were evaluated with Type II sums of squares.

We conducted separate selection analyses on data from each site, and within each site we examined selection separately for urban and rural genotypes. In addition, we asked whether the strength of selection differed between urban and rural genotypes by fitting a linear model with relative male or female fitness as the response variable, source region and the interaction between the two as predictor variables (Appendix 1: Supplementary text, A4).

We quantified stabilizing and disruptive selection using the same model format but included the square of each trait as an additional variable. When quadratic terms were significant, we fit nonparametric cubic splines to the data using the *smooth.spline* function in R to determine where there was a fitness minimum or maximum within the phenotypic range (Schluter and Nychka 1994).

We quantified selection gradients and tested whether they differ between urban and rural genotypes by fitting a linear model for the three phenology traits following the same format as the selection differential models above: relative male or female fitness as the response variable, source region, the three phenology traits, and interactions between source region and each trait (Appendix 1: Supplementary text, A5). We omitted height in the selection gradients analyses because it was highly correlated with our estimates of fitness (all $r > 0.6$).For all quadratic gradients, we multiplied the regression coefficient by two (Stinchcombe et al. 2008).

To test whether selection was stronger on foreign than local genotypes, we calculated local-foreign contrasts and conducted a one-sided Wilcoxon signed rank test on these contrasts. We calculated the absolute difference in selection differentials and gradients between local versus foreign genotypes (foreign coefficient – local coefficient) at each field site for each trait. We then used a one-sided Wilcoxon signed rank test (*wilcox.test*) to determine if the difference in selection was significantly greater than zero, which would indicate stronger selection on the foreign genotype. We also include the results of the two-sided Wilcoxon signed rank test for comparison. We conducted all selection analyses with the *lme4* package (Bates et al. 2015), and *car* packages (Fox and Weisberg 2011), and conducted separate tests for selection differentials and gradients, and for male and female fitness.

RESULTS

Urban-rural phenotypic divergence and adaptation

Data from the four sites revealed genetic differences in flowering time between urban and rural populations. Urban populations flowered earlier than rural populations ($p = 0.04$ first open male, $p = 0.0006$ first open female flower; Figure 2; Appendix 1: Table

S3), a difference that was consistent across the four field sites ($p > 0.4$ for all site × source region: Appendix 1,Table S3). By contrast, plants growing in the urban sites (USC, USP) flowered significantly later than plants in rural sites (site: first open male and female flower, $p< 0.0001$, Appendix 1: Table S3, Figure 2).

In contrast to the phenological traits, for which there were no significant site × source region interactions (Table S3), there was evidence that fitness and height of urban and rural populations differed among sites (site × source region interaction). Although these interaction terms were not significant in the GLMMs (LRT; flowers: $\chi^2 = 5.38$, $p = 0.15$, fruits: $\chi^2 = 0.89$, $p = 0.83$, height: $\chi^2 = 6.49$, $p = 0.09$), the interaction was highly significant in the aster analyses (male and female lifetime fitness: site × source region: $p < 0.0001$, Appendix 1: Table S9). This reflects the greater statistical power of aster models, obtained by including multiple fitness components and modelling each with a different statistical distribution (Geyer et al. 2007; Shaw et al. 2008).

Within each site, the GLMMs indicated there were no significant differences between urban and rural populations in either male fitness, female fitness or height (source region: $p > 0.1$ at all sites, Appendix 1: Table S4; see Table S5 for local-foreign contrasts), with the exception of the rural site RRO, where rural populations were 7 cm taller, on average, than urban populations (RRO: $p = 0.006$, Appendix 1: Table S4, Figure S3). By contrast, the aster models revealed that male and female lifetime fitness of rural populations was significantly higher than urban populations at all four sites (source region: all $p < 0.003$, Figure 3; see Appendix 1: Table S10 for local-foreign contrasts), except for USP, where urban populations had slightly higher male lifetime fitness than rural populations.

Inter-population divergence within urban and rural regions

The phenological traits, transition to reproductive phase and all flowering time traits, differed significantly among urban populations (source population: all $p<0.005$, Figure 4, Appendix 1: Table S6) and tended to differ among rural populations (source population: all $p<0.09$, Appendix 1: Table S6), with no strong evidence for the site affecting population differences (site × source population: all $p>0.2$), except for time

between first open male and female flower for rural populations ($p = 0.06$, Appendix 1: Table S6). There was no evidence for differences in the variance among urban versus among rural populations (Levene's test: $p > 0.4$ for all phenological traits).

Male fitness also differed significantly among urban populations (source population: all $p<0.05$, Appendix 1: Figure S1, Table S7) with no evidence for site affecting population differences (site × urban population: LRT, $\chi^2 = 14.69$, df = 13, $p = 0.33$). By contrast, there was evidence for a site × urban population for female fitness and height (LRT: female fitness, $\chi^2 = 22.41$, df = 13, $p = 0.05$, height, $\chi^2 = 26.65$, df = 13, $p = 0.01$), although the ranking among urban population was similar across sites (Appendix 1: Figure S2, S7). Within each site, there was a significant effect of urban population for female fitness and height (source population: all $p<0.05$, Appendix 1: Table S7).

There were no differences in either male or female fitness among rural populations (source population: $p > 0.2$ for both traits, Appendix 1: Table S8, Figure S1, S2), nor did site affect this pattern (site × rural population: $p > 0.3$ for both traits). Mean height of rural populations varied significantly among sites (site × rural population: $\chi^2 = 33.93$, df = 13, $p =0.001$), and among rural populations at the USP field site (Appendix 1: Table S8).

There was some evidence for unequal variance in fitness among urban and rural populations, however this was only true at STP (Levene'e test: male fitness, df = 6, $p = 0.001$; female fitness, df = 6, $p = 0.003$), and RRT (Levene's test: female fitness, df = 6, $p = 0.004$). This is not surprising given each site had 4-6 populations per region, and therefore there was little statistical power to detect unequal variance.

Selection analyses

Selection differentials (net selection), calculated for both urban and rural genotypes for data from each of the four experimental sites, revealed that selection tended favour larger plants, earlier transition to reproduction, earlier flowering, and delayed time between first open male and female flower (Table 1). Although not all estimated differentials were statistically significant ($p<0.05$), the estimates of selection were generally in the same direction as the patterns described above. At all field sites selection

strongly favoured large plants (Table 1). Consistent with past adaptation to local environments, assuming that local genotypes are near a fitness optimum, net selection tended to be stronger on the foreign seed source at all sites, although the pattern was not as strong for male fitness (Wilcoxon one-sided sign test: male fitness, $p = 0.07$, female fitness, $p = 0.01$; Wilcoxon two-sided sign test: male fitness, $p = 0.14$ female fitness, $p = 0.02$).

Selection gradients revealed some similar patterns for the phenological traits, although few selection gradients were statistically significant (Table 2) and there was no evidence for stronger direct selection on foreign genotypes than rural genotypes (Wilcoxon one-sided sign test of local-foreign contrasts: male fitness, $p = 0.90$, female fitness, $p = 0.38$; Wilcoxon two-sided sign test: male fitness, $p = 0.23$, female fitness, $p = 0.73$). The weaker statistical support for the gradients is not surprising given that there is some correlation among all traits (Appendix 1: Table S11, S12) and the statistical significance of each trait is evaluated after accounting for variance in fitness that can be explained by other traits. In other words, our analysis does not provide evidence for selection acting on any specific trait, but rather the cumulative effects of selection acting on a trait and those traits which are correlated with it. Although we detected significant quadratic selection differentials and gradients (Appendix 1: Table S13, S14), the non-parametric cubic splines revealed that there were no fitness minima or maxima within the range of the data. Thus, the significant quadratic parameters may reflect curvilinearity to the fitness function rather than stabilizing or disruptive selection.

DISCUSSION

To advance our understanding of adaptation to urban environments, we conducted a series of reciprocal transplant experiments, using the native annual plant *Ambrosia artemisiifolia.* Previous investigations of plant adaptation to urban environments have revealed evidence for adaptive divergence between urban and rural populations (Cheptou et al. 2008; Thompson et al. 2016; Yakub and Tiffin 2016). However, ours is the first study to examine selection and adaptation of urban populations using whole organism

phenotypes, lifetime fitness, and data collected from plants planted directly into the ground under semi-natural field conditions. Our results suggest local adaptation of both urban and rural populations: urban and rural populations have genetically diverged in multiple phenological traits – including flowering time, a trait that is often under strong selection (Hall and Willis 2006; Franks et al. 2007). Moreover, across both urban and rural sites net selection acting on phenological traits was stronger on foreign genotype, particularly with female fitness. This is consistent with local genotypes being closer to a selective optimum, as expected if populations have adapted to local conditions. However, this pattern of selection was not as strong for male fitness, nor was this result detected in the selection gradients and thus we cannot differentiate the effects of selection acting directly on our measured traits from the effects of direct selection and indirect selection acting through other measured traits.

Although the pattern of phenological divergence and the evidence of stronger selection acting on foreign genotypes are both consistent with local adaptation of urban and rural populations, the aster analyses revealed that rural populations had higher lifetime fitness than urban populations at all sites. There are several possible reasons for this apparent mismatch: selection on phenological traits may have contributed to the divergence of urban and rural populations but selection on other unmeasured traits, may be more important to lifetime fitness (Lande and Arnold 1983). Alternatively, selection and the adaptive response to selection may be stronger in the rural than in the urban environment, perhaps because the rural environment is more uniform (see discussion of spatial heterogeneity below) or because of larger effective population sizes in the rural than in the urban populations.

Counter-gradient variation

Phenotypic differences between urban and rural environments reflect both environmentally-induced, plastic effects, and genetic divergence between urban and rural populations. We found that the plants grown in the urban sites flowered later than plants in the rural sites, reflecting a plastic response that affected both urban and rural populations. However, urban populations tended to flower earlier than rural populations:

a pattern of counter-gradient variation whereby genetic differences are in the opposite direction of plastic responses (Levins 1968, 1969; Conover and Present 1990; Conover and Schultz 1995). The earlier flowering of urban populations might be due to the drier conditions and reduced water availability of urban environments compared to non-urban areas (Huff and Changnon 1973; Whitlow et al. 1992; Shepherd and Pierce 2002). Earlier reproduction to avoid drought conditions has been commonly observed, and is likely a common adaptation in annual, ruderal species such as ragweed (Grime 1977; Brachi et al. 2013; Wolfe and Tonsor 2013).

Counter-gradient variation has been found in other transplant experiments of local adaptation (Eckhart et al. 2004; Chambers and Emery 2016). Indeed, Thompson et al. (2016) reported a similar pattern in *Trifolium repens* across an urban-rural gradient: they observed the frequency of cyanogensis declined towards the urban core, but a potted transplant experiment revealed selection for increased cyanogensis in urban areas (Thompson et al. 2016). The evidence for counter-gradient variation in both our study and that of Thompson et al. (Thompson et al. 2016), the two studies that have conducted field experiments to investigate urban adaptation in plants, suggest that selection and adaptation to urban environments might often be opposite to plastic responses to environmental variation. In other words, plastic responses may not always reflect underlying genetic differences. This result underscores the importance of manipulative experiments for properly characterizing the role of selection and adaptation in driving divergence between urban and rural populations.

Our finding of earlier flowering in rural than urban field sites is opposite to ecological and herbaria data suggesting that plants growing in urban areas, including *A. artemisiifolia*, tend to flower earlier than plants growing in rural areas (Ziska et al. 2003; Neil and Wu 2006). The reasons for this are unclear. It could be related to different growth conditions, either in weather during the year we conducted our experiment, or at each of our field sites. Alternatively, such a pattern may be driven by differences in the germination timing of urban and rural populations. We exposed all seeds to the same stratification length and temperature, and then germinated them under common greenhouse conditions. However, in natural populations the germination environment and

cues of urban and rural populations may vary. For example, the urban heat island (Oke 1973) may cause plants to germinate earlier than those in rural locations. As germination timing determines the environment that subsequent life stages experience (*e.g.* flowering), it can influence the evolution and adaptation of plant populations (Donohue et al. 2010). Our results suggest exploring the potential urban-rural differences in germination and post-germination traits could be a promising direction of future research in urban plant adaptation.

Spatial heterogeneity and phenotypic differences among urban populations

Much of the experimental work in urban adaptation has treated populations collected from within the same urban area as being similar to one another (Cheptou et al. 2008; Winchell et al. 2016; Yakub and Tiffin 2016) or predicted that an "urban phenotype" will become more common in more urbanized environments (Thompson et al. 2016). Urban areas are, however, spatially heterogeneous (Cadenasso et al. 2007; Pickett et al. 2011) and this heterogeneity can be important ecologically (Knapp et al. 2012) and evolutionarily (Munshi-South and Kharchenko 2010; Lourenço et al. 2017). Consistent with this heterogeneity being evolutionarily important, we found phenological divergence and fitness differences among urban populations at a scale of 2-3 kilometers. Across all field sites, both urban and rural populations had differences in phenology among populations, but the pattern was stronger among urban populations. Furthermore, there were differences in both male and female fitness among urban populations, but not for rural populations. Thus, the urban populations appear to display greater variance in both phenological and fitness traits, although this should be viewed with caution given that unequal variances are not statistically well supported (Levene's test). However, with only a maximum of eight populations per region, there is little statistical power to detect unequal variances. In addition, the range of phenological differences among urban populations was often greater than those at the regional level, *i.e.,* among all urban and all rural populations. These results suggest that selection and adaptation may vary across spatial scales in urban environments, a pattern which has been found in other species in natural or non-urban environments (Bischoff et al. 2006; Brachi et al. 2013). By ignoring

these fine-scale differences within cities, we may be limiting our ability to explain phenotypic variation and determine which environmental variables are ultimately driving urban plant adaptation.

Our reciprocal transplant was reciprocal at the scale of urban and rural environments: we did not have field sites in the specific habitats from which we sampled each source population. Therefore, we cannot determine if there is microgeographic adaptation within either the urban or rural environment, nor which environmental variables may drive such patterns of adaptation. Furthermore, patterns of phenotypic differentiation among urban *A. artemisiifolia* populations could be driven by reduced gene flow and greater genetic drift; we cannot disentangle whether these patterns were caused by neutral processes or selection. Nonetheless, there were some notable environmental variables that varied among our collection locales and field sites. Urban collection locales varied from parks and low-density housing to highly developed areas (see Table S2). Furthermore, populations collected from sandy habitats flowered earlier and there was strong selection for earlier flowering at the field site with the highest percent sand (RRT), suggesting selection in response to sandy habitats may favour earlier flowering genotypes. Our results also support the hypothesis that mowing may lead to delayed flowering in *A. artemisiifolia* (Patracchini et al. 2011; Milakovic et al. 2014): the U8 population was sampled from an area that is mowed frequently and individuals from this population flowered later, and had lower fitness, than other urban populations. While these habitat differences alone do not indicate microgeographic adaptation, they do indicate heterogeneity within urban areas, which in turn may shape patterns of adaptation among urban populations.

In conclusion, we found evidence for genetic divergence in ecologically-important traits between urban and rural environments, but support for local adaptation was mixed. The selection analyses provided some support for selection against foreign genotypes, which is consistent with local adaptation, but rural populations overall had higher fitness at all field sites, which is not consistent with local adaptation. In addition, we found greater phenotypic divergence among urban populations than rural populations, which we hypothesize may be driven by the higher environmental heterogeneity present

in urban areas. Our results suggest that urban environments are fundamentally shaping the phenotypic evolution of plant populations. In addition, future work should consider including both multi-year experiments and fine-scale sampling to explicitly incorporate micro-environmental variation among sites and genotypes in cities.

Chapter 2

Does adaptation to historical climate shape plant responses to future rainfall patterns? A rainfall manipulation experiment with common ragweed

ABSTRACT

Climate change is affecting both the volume and distribution of precipitation, which in turn is expected to affect the growth and reproduction of plant populations. The near ubiquity of local adaptation suggests that adaptive differentiation may have important consequences for how populations are affected by and respond to changing precipitation. Here, we manipulated rainfall in a common garden to examine how differentiation among populations of common ragweed, *Ambrosia artemisiifolia* (Asteraceae) affects responses to water availability expected under climate change. We collected seeds from 26 populations along gradients of historical rainfall and used event-based rainout shelters and watering additions to simulate drier summer conditions and more extreme rainfall events, respectively. *Ambrosia artemisiifolia* had higher fitness on average under reduced rainfall, suggesting it may spread and become more abundant in areas projected to become hotter and drier during the summer months. We also found strong evidence for phenotypic and fitness clines across both latitude and longitude, and that phenological responses and fitness effects of altered rainfall depended on seed source or historical climate. The effect of rainfall treatment on female fitness was highest in western and mid longitudes, but there was little effect on eastern populations. Across latitude, the effect of rainfall treatment on male fitness was highest in southern populations. These phenology and fitness clines suggest that adaptive differentiation across the species' range has the potential to shape future responses of A. *artemisiifolia* populations to climate change, particularly altered patterns of rainfall.

Published as:

Gorton, A. J., Tiffin, P., and D. A. Moeller. 2019. Does adaptation to historical climate shape plant responses to future rainfall patterns? A rainfall manipulation experiment with common ragweed. Oecologia 190: 941-953.

INTRODUCTION

One key aspect of climate change for many organisms, including plants, is changing precipitation patterns (Trenberth et al. 2003; Weltzin et al. 2003; Wu et al. 2011). Over the next century the contiguous United States is projected to experience changes in both the overall volume and event severity of precipitation (IPCC 2013). In particular, total precipitation is projected to decline in all areas during the summer months while heavy precipitation events are projected to increase in frequency (IPCC 2013). This suggests that rain will likely fall in fewer, but more intense events during these key growing months.

Water availability is known to have consequences for plant development (Schwinning and Ehleringer 2001), physiology (Noy-Meir 1973; Dudley 1996), and species distributions (Engelbrecht et al. 2007). It follows that changes in precipitation will likely affect the growth and fecundity of individual plants and the dynamics of populations. In addition, precipitation changes may interact with rising temperatures, potentially imposing drier, drought-like conditions for plants. On the other hand, heavy precipitation events may cause episodic flooding or soil saturation, which could also impair plant-soil-water relations and plant growth and development (Kramer 1951; Kozlowski 1992). By directly affecting the growth and fecundity of individual plants, this widespread and rapid change in precipitation has the potential to affect the geographic distribution of species as well as selection on populations (Franks et al. 2007). Responses of particular species or populations to altered water availability is, however, unclear.

Manipulative experiments have shown that increased precipitation often results in greater individual above-ground biomass and ecosystem productivity while decreased precipitation tends to suppress both (Wu et al. 2011; Didiano et al. 2016). However, the magnitude and direction of responses can vary among species (*e.g.,* Campbell & Wendlandt 2013, Didiano et al. 2016). For example, Schneider et al. (2014) found that reduced watering led to a reduction in the biomass of the legume *Lupinus perennis*, but not in the C_3 grass *Agropyron repens*. Furthermore, there are mixed results as to whether geographic origin and climate influence the ecological and evolutionary responses of plants to altered precipitation patterns. A recent meta-analysis of experimental warming

and precipitation found that plant biomass and net primary productivity did not vary among geographic locations that differed in mean annual temperature or mean annual precipitation (Wu et al. 2011). By contrast, a meta-analysis of the effect of local climatic factors on phenotypic selection in wild populations found that 50-70% of the variation in selection on traits was associated with variation in precipitation (Siepielski et al. 2017). However phenotypic selection analyses cannot disentangle the effect of the environment on trait expression from that of genetics (Rausher 1992). Moreover, the environmental effects on trait expression can sometimes be opposite to patterns of genetic divergence (Conover and Schultz 1995; Eckhart et al. 2004; Gorton et al. 2018) thus phenotypic data, alone, may misrepresent expected adaptive responses to climate change.

The prevalence of local adaptation to climate (Turesson 1922; Clausen et al. 1940; Aitken et al. 2008; Leimu and Fischer 2008; Hereford 2009) indicates that adaptation has and will likely continue to play an important role in species persistence under future climates. Since the rate of contemporary climate change has occurred rapidly, population responses to selection could be constrained by limited standing genetic variation (Jump and Penuelas 2005; Hoffmann and Sgrò 2011). However, adaptation may be facilitated by gene flow, particularly if that gene flow occurs via the introduction of alleles from elsewhere in a species' range where historical climate better matches predicted future conditions (Davis and Shaw 2001; Sexton et al. 2011; Aitken and Whitlock 2012). Nonetheless, the extent to which potential gene flow might facilitate adaptation to future environments is dependent on the extent to which populations are adaptively differentiated.

Experimental manipulations of precipitation are required to disentangle evolutionary adaptation from phenotypic plasticity. Long-term observation of natural populations in response to climate can provide important insight into the ecological effects of ongoing precipitation changes on plant traits (*e.g.,* Dunnett et al. 1998). However, observational data do not bear on adaptive differentiation because they inherently confound phenotypic plasticity with genetic differentiation. Furthermore, altered precipitation may be confounded with other covarying factors, such as temperature or CO_2 levels. Thus, experimental manipulations of precipitation in a

common garden environment importantly complement observations of natural populations (Beier et al. 2012). While reciprocal transplant experiments are required to test for local adaptation, common garden experiments can allow patterns of genetic differentiation to be identified, and in turn, correlated with associated environmental differences at each collection site. Common garden experiments also offer greater control of confounding variables, allow multiple climate change scenarios to be studied simultaneously, and can more readily allow researchers to disentangle the contribution of the environment vs genetics to trait expression and their consequences for plant population responses to climate change.

Here, we manipulated rainfall onto plots in a common garden to ask how simulated changes in rainfall affect trait expression and lifetime fitness among populations of common ragweed, *Ambrosia artemisiifolia* (Asteraceae) and whether those responses varied relative to the historical climate experienced by those populations. Our experimental plants were grown in a common garden from seeds that we collected from source populations sampled along both latitudinal and longitudinal gradients as well as gradients of historical annual rainfall. We manipulated water availability through the use of event-based rainout shelters to reduce total rainfall and simulate drier summer conditions, and watering additions to increase total rainfall and simulate more extreme rainfall events. We collected data on ecologically-important traits (flowering time, size, specific leaf area), and components of fitness (flower and fruit number). We address the following specific questions: (1) How will future rainfall patterns affect the growth, phenology and fitness of *A. artemisiifolia*? and (2) What is the nature of phenotypic differentiation in response to historical climate variation along latitudinal and longitudinal gradients? and (3) To what extent does historical rainfall at collection origin predictive of the response to altered water availability, *i.e.*, is the response dependent on the latitude and longitude or annual precipitation of each seed source?

METHODS

We conducted a rainfall manipulation experiment in Minnesota, in the northern portion of the species' range, using seeds sampled from populations across a large portion

of the geographic range of *Ambrosia artemisiifolia* (Asteraceae). In the midwestern United States, the summer growing months are projected to become both hotter and drier on average, with a reduction of 10-30% in seasonal precipitation depending on the location (IPCC 2013; NOAA 2014). Furthermore, in the past 50 years, there has been a 37% increase in the amount of precipitation falling in heavy rainfall events (NOAA 2014, pg 9). Based on these historical trends and projections, we chose to reduce rainfall by 30% and increase rainfall by 30% for the reduction and addition treatments, respectively (see *Rainfall manipulation experiment* below).

Ambrosia artemisiifolia L. (Asteraceae) is a self-incompatible (Friedman and Barrett 2008), monoecious, and wind-pollinated (Jones 1936; Essl et al. 2015) annual plant native to North America and invasive on multiple continents (Bass et al. 2000; Chauvel et al. 2006). It is a ruderal plant that is often abundant in open, disturbed habitats such as river banks, roadsides, agricultural fields, and urban areas. *Ambrosia artemisiifolia* is a summer annual that typically germinates in late spring in Minnesota (May-June) and flowers in late summer. It is sensitive to freezing and can be killed by late spring and early autumn frosts; the latter terminates the growing season.

The transition to reproduction is cued by photoperiod and is initiated when the length of day shortens sufficiently after the summer solstice. Staminate capitula (*i.e.* male flowering heads) are found in spike-like racemes, hereafter referred to as 'male flowers', which produce the pollen that is one of the primary causes of summer and fall allergic rhinitis (Lewis et al. 1983; Frenz 2001). Pistillate capitula (*i.e.* female flowering heads) are found in axillary clusters below the male flowers; each individual flower develops into an achene (a small, single seeded fruit) which readily falls off the plant once ripe. These groups of achenes are hereafter referred to as 'fruits'.

Seed collections

During Oct 2015 – Jan 2016, we collected seeds from 26 populations of *A. artemisiifolia* across a region spanning 15 degrees of latitude (~1700 km) and 7 degrees of longitude (~550 km). These populations have experienced a wide range of historical combinations of temperature and precipitation (Appendix 2: Figure S1, Table S1). The

sampling area includes both a north-south and an east-west precipitation gradient that ranges from a low of 609 mm per year in eastern South Dakota to a high of 1595 mm per year in southern Louisiana (precipitation data from WorldClim 2.0). Sampling sites were on average 290 km apart along each latitudinal transect, and when two populations were sampled at a given latitude, we collected separate populations in both urban and rural environments because the phenotypes of urban and rural populations of *A. artemisiifolia* differ from those of rural populations (Gorton et al. 2018). At each sampling site, we collected seeds from 16-25 maternal plants, each separated from the rest by at least 3 meters.

Rainfall manipulation experiment

In May 2016, we planted a single common garden on the University of Minnesota campus in Minneapolis, Minnesota, in a previously abandoned field (44.9763, -93.21948) that was densely occupied by weeds. Prior to planting seeds, we sprayed the area with glyphosate and tilled with a tractor to remove existing vegetation, creating an environment similar to the disturbed, low-competition habitats where ragweed most often occurs.

We established 30 rainfall treatment plots (10 reduction, 10 addition, 10 control) in a 9 column x 3-4 row grid with 1 m spacing on all sides around each plot (Figure S2). The planting spacing within each plot was designed to mimic natural growing densities of *A. artemisiifolia* (Foster et al. 1980, MacDonald and Kotanen 2010). The plots alternated across the grid to ensure the rainfall treatments were distributed evenly across the study area. The 3-4 plots in each column constituted a 'block' (Appendix 2: Figure S2). We chose blocks in this way because the columns occurred perpendicular to a gradual slope that had the potential to influence rainfall runoff and soil moisture. Each plot measured 2 m^2 and included a 40 cm border from the edge of the experimental plants to the edge of the plot. Each plot contained seeds from one to two different families from each of the 26 populations, for a total of 45 seeds per plot, planted in a completely randomized design. A total of 16-25 families per population were included in the experiment (total number of families = 367) (Appendix 2: Table S1).

A. artemisiifolia has strong dormancy, with viable seeds remaining in the seed bank for up to 40 years (Toole and Brown 1946). To minimize the potential to confuse experimental plants with plants that grew from seeds in the seed bank, we used the ProPlugger (ProPlugger, NC) to remove a soil plug (5 cm x 5.5 cm) at each planting spot. The holes were then filled with B2 germination mix (Berger, Quebec). The seeds were stratified in moist silica sand to break dormancy, and kept in the dark at 4°C for 10 weeks (Willemsen 1975). At each planting spot, we planted 2-4 stratified seeds from a given maternal family directly into the ground from 31 May to 3 June 2016, and watered the seeds once immediately after planting (total number of seeds planted = 4515). Plants in the reduction and control plots were not watered again for the duration of the experiment. Two weeks after planting, plants were thinned to a single seedling per planting spot. Throughout the growing season, all plots were regularly weeded to minimize interspecific competition.

We used event-based rainout shelters for our reduction treatments. These types of shelters are removable and are deployed to exclude specific rainfall events rather than remain in place for the duration of the growing season (*e.g.* Eisenhauer et al. 2012, Reich et al. 2014). Consequently, they minimize microhabitat effects on the underlying vegetation, including shading, passive warming, and altered humidity and wind (Beier et al. 2012). We designed and constructed the rainout shelters using PVC pipes and clear, overwintering greenhouse plastic (Appendix 2: Figure S3). Beneath the lowest point of the roof of each reduction plot, we placed a rain barrel to collect excluded rainfall. We also placed a rain gauge in the center of each reduction plot to determine if the rainout shelters failed during an excluded rainfall event.

We conducted the rainfall manipulation treatments by attaching the roofs of the rainout shelters for 11 rainfall events from 30 June 2016 until 20 September 2016. We removed the rainout shelter roofs the day after a rainfall event, and recorded the total rainfall in uncovered rain gauges. We calculated the total reduction in rainfall from June-September as the ratio of the summed recorded rainfall in the rain gauges on the days we deployed the rainout shelters to the total rainfall for the season, obtained as the sum of the

amounts in the rainfall gauges over the whole season. Based on these data, we excluded approximately 30% of total rainfall from the exclusion plots.

For the addition plots, we added supplementary water the day after an excluded rainfall event (n = 11) to increase rainfall by 30%. We multiplied the rainfall depth recorded in the uncovered rain gauges by the area occupied by plants in each addition plot (1.69 m^2) to calculate the volume of water to be added to each addition plot. We used water from the rain barrels, when available, and applied it to the addition plots using watering cans to evenly distribute the water. Early in the season when rain barrels were empty, we used tap water for the addition treatments.

Phenotypic and environmental data collection

Each week we recorded whether each plant was in a vegetative or reproductive phase and whether they had produced their first male and female flower. The transition to reproduction was scored as the date of the first appearance of a reproductive bud at the apical meristem. The first male flower was scored as the date on which the first anther opened and shed pollen, *i.e.* the first open male flower, and the first female flower was scored as the date at which stigmas first appeared. We also measured plant height at 8 and 19 weeks after planting.

We estimated specific leaf area (SLA), the ratio of leaf area/leaf dry mass (cm^3/g) as a proxy for water use efficiency (Reich et al. 1991; Poorter and Bongers 2006). We estimated SLA using the first fully-expanded leaf from each plant eight weeks after planting, approximately four weeks after the start of the rainfall treatments. We stored all leaves at 4°C in individual ziplock bags with a moist paper towel to prevent wilting, and scanned them within three days of harvest with a leaf area meter (LI-Cor LI-3000A Portable Area Meter and LI-3050A Transparent Belt Conveyor), followed by drying at 55°C for 7 days. Once the leaves were completely dry, we weighed them to obtain dry mass.

We estimated fitness at the end of the growing season between 24 September and 11 October, which was after individuals stopped growing and when plants begin to senesce. Based on data from 1981-2010, there is a 50% chance of a frost occurring by 5

October (0°C, U.S. Climate Normals, NOAA). It was not possible to quantify male and female fitness across whole plants due to their size and the narrow time window over which data collection needed to occur. Instead, we systematically subsampled branches across each plant to account for potential variation in allocation to male versus female reproduction across plant development (ca. 20-50 total branches/plant). On each plant, we counted the number of flowers and fruits on every fourth branch, starting with the lowest and largest branch, and moving up towards the apical meristem. We multiplied this number by four to get a whole-plant estimate of male and female fitness.

We collected data on soil moisture (volumetric water content, VWC) every hour from the time seeds were planted to the time of harvesting, using ECH_20-E5 Decagon soil moisture sensors and Em50 ECH_20 data loggers. We buried soil moisture probes in the center of 10 treatment plots and inserted them horizontally at a depth of 30 cm (4 reduction plots, 4 addition plots, and 2 control plots). Due to probe malfunctioning (negative VWC values and/or incorrect dates), we were unable to obtain data from the beginning of the rainfall manipulation treatments; usable data began around 29 July 2016, approximately one month after the start of the rainfall treatments. Data from the soil moisture probes indicated the reduction plots had drier soil than the control and addition plots (Figure S4), indicating the efficacy of our treatments in altering water availability.

We extracted historical precipitation data (1970-2000) based on the latitude and longitude of each collection site from WorldClim 2.0 (Fick and Hijmans 2017). We downloaded the BIOCLIM precipitation variables (BIO12-BIO19) for each location at a spatial resolution of 5 m.

Statistical analysis

To determine how phenology (date of transition to reproductive phase, date of first open male flower, date of first open female flower), height (at 8 and 9 weeks after planting) and SLA responded to the rainfall treatments, we fit linear models (LMs, *lme4* package, Bates et al. 2015) with rainfall treatment as the explanatory variable of interest. In these models we also included block (column in which each plot occurred) and edge (whether the individual plant was on the edge or center of the treatment plot) as

categorical variables to account for effects of microenvironmental variation across the field site (Figure S2). We ran separate models with each of the three phenology traits (date of transition to reproductive phase, date of first male flower, date of first female flower), each of the two height measurements (height at 8 weeks, height at 19 weeks), and SLA as the dependent variables. We tested for differences among rainfall treatments using Tukey's tests implemented in the *multcomp* package (Hothorn et al. 2008). In addition, since we conducted the same analysis on six phenotypic traits, we adjusted the alpha level required to infer significance for each test using a sequential Holm-Bonferroni correction (Holm 1979) using *p.adjust*() in base R.

Many climate variables, including annual precipitation, annual temperature, and growing season length are highly correlated with one another as well as with latitude or longitude. Consequently, to examine the effect of seed source on the response to rainfall treatment, we conducted two sets of analyses with different environmental predictors. The first used latitude and longitude as predictors, considering them proxies for historical climate and other geographically variable environmental factors. Due to the low range of longitudes included across the sampled populations, we binned longitude into three categories for all analyses: 1, longitudes greater than 94° ('western'), 2, longitudes between 92° and 94° ('central), and 3, longitudes less than 92° ('eastern'). In the second set of analyses, for each collection site we used the value for the first principal component (PC1) of the seven BIOCLIM precipitation variables as a predictor (PC1 accounted for 85% of the variation, Table S2, S3). PCs were obtained using *princomp*() in base R. To determine whether phenology, height and SLA varied with latitude and longitude of seed source and rainfall treatment, we fit LMs with block (column in which each plot occurred), edge (whether plant was located on the edge or the middle of plot), latitude and longitude of origin of each population, rainfall treatment, latitude × longitude, latitude × treatment, and longitude × treatment as predictors. We ran similar models with the precipitation PC1, instead of latitude and longitude, as a predictor. We also tested for non-linearity in the relationship between each trait and latitude by including latitude^2. In cases where the quadratic term was significant, we fit splines using *smooth.spline()* to examine the shape of the relationship. For each of the phenology and

size traits, latitude^2 explained a small proportion of the variance, and was statistically significant only for height at 8 weeks (Appendix 2: Table S4). There was no evidence of any major curvature nor a plateau in our data (Appendix 2: Figure S5), therefore we have presented only the linear analyses in the results below.

In both sets of analyses, as before, we tested for differences among rainfall treatments using Tukey's tests and controlled for multiple testing using a sequential Holm-Bonferroni correction (Holm 1979). For all LM analyses, we determined the significance of the predictors and interactions using F-tests and estimated percent variance explained using Type II sums of squares with the *car* package (Fox and Weisberg 2011). Least-square means were extracted for plots from each model using the *emmeans* package (Lenth 2016).

To determine how lifetime fitness was affected by rainfall treatment and whether it varied among seed sources, we used fixed effects models in '*aster*' (*aster* package, Geyer et al. 2007, Shaw et al. 2008). *Aster* explicitly models the dependence of fitness components expressed later in development (*e.g.* fecundity) on those expressed earlier (*e.g.* survival) and allows for different statistical distributions for each fitness component. For each individual, we used the following graphical model to estimate lifetime female fitness (number of fruits):

$$\mathbf{1} \rightarrow \underset{(0,1)\ Bernoulli}{\boldsymbol{Surivval\ to\ 8\ weeks}} \rightarrow \underset{(0,1)\ Bernoulli}{\boldsymbol{Any\ Fruit\ or\ Flowers}}$$

$$\rightarrow \underset{zero-truncated\ Poisson}{\boldsymbol{Number\ of\ fruit}} \rightarrow$$

The same model was used for male fitness, except that the number of male flowers replaced number of fruits. In addition to the response variable of lifetime fitness, we included the same predictors as above (block, edge, latitude, longitude, latitude × longitude, latitude × treatment, longitude × treatment). We ran similar models with the precipitation PC1, substituting PC1 for latitude and longitude. As in our LMs, we also tested for non-linearity between fitness and latitude by including latitude^2, and fit splines using *smooth.spline()* to examine the shape of the relationship. As there was evidence for a significant effect of latitude^2 on both male and female fitness, we also tested whether this relationship changed depending on the rainfall treatment by including latitude^2 ×

treatment. We determined the significance of all fixed effects and interactions by sequentially comparing nested models with and without the term of interest using likelihood ratio tests. For significant interactions, we predicted male and female lifetime fitness separately for each combination. Hereafter, predictions of lifetime fitness from the aster models are referred to as 'fitness'.

To better understand how individual components of fitness responded to treatments and jointly contributed to lifetime fitness, we ran separate GLMs for each of the three fitness components: survival to 8 weeks, probability of fruiting (female) or flowering (male), and number of fruits or flowers. We used a binomial error distribution and logit link function for survival to 8 weeks and probability of flowering/fruiting, and a Poisson error distribution and log link function for number of fruits and number of flowers. Unlike the analyses we ran in *aster*, these analyses are conditional such that individuals are only included in the analysis if they successfully reached that life history stage (e.g. only plants that survived to 8 weeks are included in the analyses for probability of fruits). In addition to treatment, these models also included edge and block as predictors. We determined the significance of each predictor using F-tests and estimated percent variance explained using Type II sums of squares with the *car* package (Fox and Weisberg 2011). Least-square means were extracted for each fitness component using the *emmeans* package (Lenth 2016). All analyses were conducted in R, 3.2.2 (R Core Team 2013).

RESULTS

Effect of rainfall treatment on traits and fitness

We detected no overall effect of rainfall treatment on either the date at which plants transitioned to reproduction or flowered ($p > 0.5$ for all phenology traits, Appendix 2: Table S5; Figure 1a). Plants in the reduction treatment were on average 4 cm shorter than those in the other treatment plots after eight weeks of growth ($p < 0.001$; Appendix 2: Table S5; Figure 1b), but this effect disappeared by the end of the growing season ($p > 0.9$; height at 19 weeks, Appendix 2: Table S5). Rainfall treatment also had an effect on specific leaf area (SLA) ($p < 0.05$, SLA, Appendix 2: Table S5; Figure 1c); plants in the

addition plots had significantly higher SLA (thinner leaves) than those in the control plots (Tukey contrast: $p = 0.003$). Plants in the addition plots also had higher SLA than those in the reduction plot, but this difference was not statistically significant (Tukey contrasts: $p = 0.11$). Fitness differed strongly among treatments: plants in the reduction plots had 31-38% higher male and 16-25% higher female fitness than plants in the control and addition plots, respectively ($p = 0.04$ for female fitness, $p < 0.0001$ for male fitness, Appendix 2: Table S6; Figure 2a,b). Plants in the control plots also had 4% higher male and 8% higher female fitness than those in the addition plots.

In our conditional analysis of each life history component in the *aster* model, we found a significant effect of treatment on survivorship at 8 weeks ($p_{\text{Wald } x2= 9.03,\ df = 2} = 0.01$), number of fruit ($p_{\text{Wald } x2= 1592.3,\ df = 2} < 0.0001$), and number of male flowers ($p_{\text{Wald } x2= 2280.3,\ df = 2} < 0.0001$). Mean survivorship at 8 weeks in the addition, control and reduction treatment was 57.3%, 68.5% and 64.1%, respectively. The mean number of fruits in the addition, control and reduction treatment was 740, 665, and 761 fruits, respectively, and the mean number of male flowers in the addition, control and reduction treatment was 192, 161, and 224, respectively. We did not detect an overall effect of treatment on either probability of fruiting ($p_{\text{Wald } x2 = 1.5,\ df = 2} = 0.47$, addition = 90.5%, control = 90.3%, reduction = 92.8%) or flowering ($p_{\text{Wald } x2= 0.41,\ df = 2} = 0.81$, addition = 84.2%, control = 86.1%, reduction = 84.3%).

Interactive effects of rainfall treatment and geographic origin on traits

Flowering time varied among populations along both latitudinal and longitudinal clines. Averaged across treatments, populations from northern latitudes and eastern longitudes initiated reproduction and flowered earlier than populations from southern latitudes and western longitudes (latitude: $p < 0.0001$; longitude: $p< 0.05$ for all phenology traits, Appendix 2: Table S7). The relationships between transition of reproduction, time to first open male flower and geographic origin were also affected by rainfall treatment (latitude × treatment: $p = 0.005$ for transition of reproduction, $p = 0.02$ for time to first open male flower, Appendix 2: Table S7), although the p-value for time to first open male flower was > 0.05 after controlling for multiple testing. In general,

plant populations from more southern latitudes initiated reproduction and flowered earlier in response to the rainfall reduction, whereas those from northern latitudes flowered at similar times across all three treatments (Appendix 2: Figure S6).

Size at both 8 and 19 weeks also displayed latitudinal clines ($p < 0.01$, Appendix 2: Table S7). Populations from southern latitudes tended to be taller, and this effect was more prominent at the end of the growing season (Appendix 2: Figure S7). SLA was lower in western than in eastern populations ($p = 0.006$, Appendix 2: Table S7, Figure S8), but there was no significant difference among the longitude categories (Tukey's test: $p > 0.05$ for all comparisons) or with latitude ($p = 0.07$, Appendix 2: Table S7).

Interactive effects of rainfall treatment and geographic origin on fitness

In general, plants in the reduction plots had the highest female fitness, but the magnitude of difference among rainfall treatments depended on the source longitude (treatment × longitude: $p = 0.03$, Appendix 2: Table S8; Figure 3). Populations from eastern longitudes had similar female fitness in all rainfall treatments, while those from central and western longitudes had the highest female fitness in the reduction plots and the lowest fitness in the addition plots. Female fitness was also related to the latitude of the source population and this relationship was non-linear (latitude^2: $p < 0.0001$, Appendix 2: Table S8) that was largely driven by the lower fitness of southern populations (Appendix 2: Figure S9a) and was not affected by treatment (latitude^2 × treatment: $p = 0.33$, deviance = 2.24).

Similar to female fitness, plants in the reduction plots had the highest male fitness, with rainfall treatments having the greatest effect on populations from southern latitudes (treatment × latitude: $p = 0.0004$, Appendix 2: Table S8; Figure 4). Male fitness of populations from southern latitudes (< 35 degrees), was approximately 5 to 9 times higher in the reduction plots compared to those in the addition or control plots. There was evidence for non-linearity in the relationship between male fitness and latitude (latitude^2: $p < 0.0001$, Appendix 2: Table S8) that was again likely driven by the lower fitness of southern populations (Appendix 2: Figure S9b). Unlike female fitness, however, the non-linear relationship between male fitness and latitude varied with treatment (latitude^2 ×

treatment: $p = 0.001$, Appendix 2: Table S8), whereby individuals from the most southern populations had much higher fitness in the reduction plots than in addition and control plots (Appendix 2: Figure S9b).

Both female and male fitness varied with both the latitude and longitude of source populations, and these effects were not uniform across either latitude or longitude (latitude × longitude: $p < 0.0001$, Appendix 2: Table S8, Figure S10). Populations from higher latitudes and western longitudes had lowest female fitness whereas populations from eastern longitudes and mid latitudes had the highest female fitness. Populations from central longitudes had similar female fitness across all latitude of origin. The patterns were similar for male fitness, except that populations from higher latitudes and central longitudes had higher male fitness than those from eastern and western longitudes.

Interactive effects of rainfall treatment and source climate on traits and fitness

The effects of the precipitation variables, as summarized by PC1, were similar to the effects estimated for latitude and longitude. On average across treatments, populations from locales with historically higher precipitation flowered later (PC1: $p < 0.0001$, for all phenology traits, Appendix 2: Table S9). The relationships between transition of reproduction, time to first open male flower and source climate were also affected by rainfall treatment. (PC1 × treatment: $p < 0.03$ for transition of reproduction and first open male flower, Appendix 2: Table S9, Figure S10), although the probability of this occurring by chance was >0.05 after controlling for multiple tests. Populations from wetter climates tended to transition to reproduction and flower earlier in response to the rainfall reduction than populations from drier climates, which transitioned to reproduction later under reduced rainfall. Populations from wetter climates also tended to be taller at the end of season (PC1: $p < 0.0001$ for height at 19 weeks, Appendix 2: Table S9). SLA was not associated with source climate (PC1: $p = 0.65$; Appendix 2: Table S9).

On average, populations from wetter climates had lower male and female fitness relative to those from drier climates (PC1: $p < 0.0001$, Appendix 2: Table S10; Figure 5a,b). There was a significant interaction between source climate and rainfall treatment

for male fitness (PC1 × treatment: p = 0.02, Appendix 2: Table S10) but not female fitness (PC1 × treatment: p = 0.44, Appendix 2: Table S10). However, there was no major difference in slope among treatments for male fitness (Figure 5b).

DISCUSSION

Differentiation across a species' range might have important consequences for how populations are affected by ongoing climate change and their potential for range shifts. One might expect that populations that have experienced histories similar to future climates will be less affected. Gene flow among populations might also mitigate effects of climate change on fitness by introducing adaptive genotypes or alleles from populations that have adapted to similar climates present elsewhere in the range. The effects of local adaptation have begun to be integrated into species distribution models (Hällfors et al. 2016, Peterson et al. 2018) and empirical work (Aiken et al. 2008, Wadgymar et al. 2018) that predict responses to climate change. However, we are not aware of any empirical experiments that have incorporated local adaptation to historical precipitation to predict plant responses to future rainfall patterns.

The results from our manipulative experiments provide some support for the hypothesis that adaptive differentiation may shape population responses to future rainfall environments. We found that ragweed populations generally had higher fitness under reduced rainfall, and we found strong evidence for phenotypic clines across both latitude and longitude. Although the overall effects of rainfall treatment and seed sources were much stronger than the interaction between the two, the phenological responses to and fitness effects of altered rainfall varied among populations in relation to their historical climate. Populations from western longitudes with historically drier climates tended to have higher fitness in the reduction treatment and also had the largest SLA, suggesting adaptive differentiation in response to precipitation. However, populations from southern latitudes, which have historically wetter climates, also had highest male fitness under reduced rainfall. While these results appear to be in conflict, they are most likely due to the interaction between latitude and longitude: western longitudes are consistently dry across latitude, whereas eastern longitudes are only dry at northern latitudes. In addition,

there are other environmental variables (e.g. temperature and photoperiod) that are changing across latitude and longitude which likely influence plant fitness. Nonetheless, the correlations between historical rainfall and performance under manipulated rainfall suggest that adaptive differentiation may determine the response of *A. artemisiifolia* to the direct effects of altered rainfall under climate change. Given this geographic variation, it follows that gene flow among populations has the potential to introduce alleles that confer higher performance under projected drier conditions during summer months.

Phenotypic responses to rainfall manipulation

Manipulative experiments have found that increased rainfall generally increases plant growth and reproduction, while reduced rainfall limits it (Wu et al. 2011; Didiano et al. 2016). We found the opposite pattern: *A. artemisiifolia* had the highest fitness in the reduction treatment, and lowest fitness in the addition treatment. This effect was not caused by greater interspecific competition in the increased rainfall treatment as we prevented differences in surrounding plant growth across treatments by weeding. Plants in the reduced rainfall treatment also had thicker leaves, suggesting reduced water availability may lead to lower SLA values.

Overall, these results suggest that populations in the portion of the species' range that currently experience the highest rainfall (mid-central southern United States), may increase in size in response to the reduced rainfall that is predicted over the next 100 years (IPCC 2013). This portion of the species' range (including AR, OK, KS, TX and LA) lies at the southern extent of the range where populations are considerably less common than in other parts of the range of *A. artemsiifolia* (Kartesz 2013). At the northern portion of its range, summer rainfall also is projected to decrease under a high emissions scenario although much less dramatically than in the south (IPCC 2013). Therefore, *A. artemisiifolia* may expand its range in the northern portion as well. Unlike the responses of many species to climate change, these results suggest *A. artemisiifolia* may expand its range and become more abundant, in particular at the southern portion of its range due to changing precipitation patterns. Nonetheless, our conclusions should be

viewed with some caution given that a portion of the summer in which we conducted the experiment was particularly wet (the sixth wettest August on record for Minneapolis). Indeed, the high volume of ambient rainfall may provide one explanation for the reduced female fitness under the addition treatment, at least for populations from central and western longitudes which may be better adapted to drier environments than the populations sampled from eastern longitudes. As we discuss below, experiments over multiple years are needed in order capture the effects of extreme events on plant growth and reproduction.

Latitudinal and longitudinal clines of phenotypic differentiation

In our common garden, we found strong evidence for among-population trait divergence across latitude and longitude. Flowering time and height varied across both latitude and longitude of sources and these clines were associated with fitness: northern and western populations flowered earlier, were shorter, and had higher fitness than southern and eastern populations. There also was a longitudinal cline in SLA, whereby western populations had thicker leaves than eastern populations. Given that flowering time, size and SLA are often ecologically-important traits in herbaceous plants, these clines are suggestive of local adaptation to environmental variation across the native range. However, neither latitude nor longitude impose selection. Rather these are proxies that serve to capture geographic variation in temperature, water availability and growing season length, and other environmental factors. Although it is likely a combination of environmental variables has caused these phenotypic clines, our experiment cannot determine which environmental factors are the causative drivers of selection, assuming that the differences are in fact adaptive.

The phenotypic divergence we found may have important consequences for range shifts under climate change. Our results indicate that there is widespread phenotypic differentiation across the species' range, suggesting that ragweed may readily adapt to new climatic conditions either *in situ* or at the expanding range edge. Furthermore, gene flow could facilitate adaptation, especially in plants with high dispersal capacity such as those with wind-dispersed pollen or seeds. Indeed *A. artemisiifolia* has historically

experienced large and rapid changes in its distribution corresponding with Pleistocene climate change and subsequent human activity (Martin et al. 2014). Our results and these historical distribution patterns suggest that *A. artemisiifolia*, and perhaps other species with widespread distributions and high dispersal capacity, may have high range-shifting potential due to the ability to migrate, to adapt to new environments, and gene flow.

Latitudinal clines in ecologically important traits are commonly found in widely distributed, temperate plant species (*e.g., Arabidopsis*: Agren & Schemske 2012; *Poplar*: Keller et al. 2011, and have been previously documented in *A. artemisiifolia* in North America (Hodgins and Rieseberg 2011), Europe (Leiblein-Wild and Tackenberg 2014) and China (Li et al. 2014). These clines likely reflect adaptation to growing season length, summer temperatures and water availability. Longitudinal clines like those we found are less often reported. Given that daylength does not change across longitude, these clines in ecologically-important traits (*e.g.* flowering time, SLA, etc) may offer an opportunity to disentangle which environmental variables underlie adaptive differentiation (Samis et al. 2008, 2012). Furthermore, the interaction between latitude and longitude for fitness indicates that migration due to climate change might not be as simple as plant distributions shifting northwards as the climate warms. Instead, the east-west migration of populations or alleles across longitude may be equally important to the potential for range shifts and adaptation of plant populations in response to changing environmental conditions.

The effect of adaptive differentiation on plant responses to climate change

If local adaptation strongly affects plant phenological responses and fitness under climate change, we might expect populations to vary in their response to the rainfall treatments based on their geographic origin, and as a consequence, vary in their ability to cope with changing precipitation patterns in the future. However, the consequences of intraspecific variation for fitness and phenology under simulated rainfall conditions are still largely unknown (Liancourt et al. 2013). Those which have incorporated intraspecific variation into experimental manipulations of rainfall have either focused on functional traits in long-lived species (*e.g., Quercus*: Cavender-Bares & Ramírez-

Valiente 2017, Ramírez-Valiente & Cavender-Bares 2017, Ramírez-Valiente et al. 2018; *Fagus*: Baudis et al. 2014) or examined biomass in response to both temperature and precipitation (Beierkuhnlein et al. 2011). By contrast, we are not aware of studies of lifetime fitness in response to rainfall manipulation.

In our experiment, there were some intriguing patterns that suggest adaptive differentiation may be important to predicting species' responses to future rainfall conditions. Populations from southern climates flowered earlier under reduced rainfall while populations from northern climates flowered later. Although we cannot determine what processes underly these patterns nor can we unequivocally state they are adaptive, one hypothesis is that populations from different climates have alternate plastic responses to drier conditions. Populations from southern latitudes or historically wetter environments may respond to dry conditions by flowering rapidly, while those from northern latitudes may tolerate dry conditions and flower later. We cannot exclude the potential for maternal environmental effects to have contributed to the patterns we found, as we used field collected seeds. However, maternal environmental effects are more likely to influence an individual's phenotype during earlier life stages (reviewed in Roach & Wulff 1987), and have not been previously found in *A. artemisiifolia* (Hodgins & Rieseberg 2011). We also found that populations from western climates had higher SLA across all treatments and higher female fitness in the reduction treatment, while populations from eastern longitudes had lower SLA across all treatments and similar fitness across treatments. Given that western populations experience lower annual precipitation, these results suggest potential adaptive differentiation in response to water availability and indicate that local adaptation may contribute to variation in the ability of different populations to respond to changing rainfall patterns. To develop a more complete understanding of how adaptive variation among populations might affect responses to climate change, it would be beneficial to extend the experiment we conducted over multiple years. This would assist in determining the extent to which population differences are affected by the absolute amount of rainfall as well as interacting factors such as temperature (Wu et al. 2011; Beier et al. 2012). Nevertheless, our results provide new information on how rainfall regimes affect lifetime fitness of

genotypes from across a species' range and thereby influence the potential for range shifts.

Chapter 3

Population genomics of local adaptation in *A. artemisiifolia*

ABSTRACT

Population genomic analyses can be used identify candidate loci which may be involved in local adaptation by testing for associations between allele frequencies and the environmental variables that are hypothesized to drive selection. We combined RNA-seq, environmental, and phenological data to examine the population genomics of local adaptation in the native range of common ragweed, *Ambrosia artemisiifolia.* We sequenced 74 individuals from 15 populations across a latitudinal range spanning Minnesota to Louisiana. With these data, we assembled a reference transcriptome *de novo,* and identified over 300,000 SNPs. Consistent with previous studies of this species, we found a pattern of isolation by distance, but the pairwise F_{ST} estimates indicated that there is weak population structure across the sampled range. We next identified candidate loci which may be involved in local adaptation using a genetic-environment association method, LFMM2. We detected candidate loci involved in a variety of functions, including flowering time, abiotic stress tolerance, seed dormancy and defense against pathogens. These loci suggest future directions for experiments investigating local adaptation in common ragweed.

Note on contributions:

The work of this chapter was done in collaboration with Tuomas Hämälä. I conceived of the questions addressed in this chapter, collected the samples, completed the RNA extractions, interpreted the results, and wrote the majority of the chapter. Tuomas Hämälä completed all of the data processing and population genetic analyses (including assembling the transcriptome, estimating population structure and diversity, running the LFMM2 analyses) and wrote the associated methods portions of the chapter.

INTRODUCTION

Local adaptation is generated by geographically variable selection. Traditionally, organismal based approaches like reciprocal transplant experiments have been used to identify locally adapted traits and estimate selection on those traits (Clausen et al. 1940; Kawecki and Ebert 2004; Geber and Eckhart 2005; Agren and Schemske 2012). Such experimental approaches do not account for historical population structure, which may contribute to our understanding of both phenotypic and genetic differentiation. Population genomic analyses provide a complementary approach to examine local adaptation by testing for associations between allele frequencies and environmental differences while also providing insight into demographic history and population structure (Tiffin and Ross-Ibarra 2014; Rellstab et al. 2015). In turn, phenotypic data from common garden experiments can inform the selection of biologically relevant environmental variables and contribute to the subsequent interpretation of signals from population genetic data (e.g., Fournier-Level et al. 2011; Yoder et al. 2014). In combination, population genomic, environmental, and phenotypic data can provide insight into both the genomic basis of adaptation and the environmental variables that may be driving divergence.

Across geographic space, migration and genetic drift lead to a pattern of isolation by distance (IBD), whereby genetic differentiation among populations increases with increasing geographic distance (Wright 1943). While the majority of loci are expected to display a pattern of IBD, portions of the genome that contribute to local adaptation are expected to have divergence patterns that differ from the genomic background, assuming the environment does not closely covary with geographic distance. These loci that may be involved in local adaptation are expected to show a pattern of isolation by environment (IBE), although identifying the important environmental variables that drive local adaptation may be challenging (Shafer and Wolf 2013; Wang et al. 2013). There are many methods available to identify the genomic targets of local adaptation, including outlier tests, genetic-environment association, QTL mapping, and GWAS (De Mita et al. 2013; Jones et al. 2013; Hoban et al. 2016). The focus of this chapter is on genetic-environment association methods or GEA methods. GEA methods identify candidate loci

that may be involved in local adaptation based on associations between genomic data and the environmental variables that are hypothesized to drive selection (Rellstab et al. 2015; Forester et al. 2018).

One challenge of identifying the loci involved in local adaptation is that geographic distance and environmental distance often covary (Bradburd et al. 2013; Hoban et al. 2016; Rellstab et al. 2016). Populations that are far apart geographically are more likely to undergo different environmental conditions than two populations that are close together, depending on the scale of variation in the environment. Similarly, populations that are closer together tend to be more related to one another than to populations originating from further locations. As such, determining whether patterns of divergence among populations might be better explained by neutral processes or selection in response to environmental differences can be difficult. Furthermore, certain demographic scenarios can be particularly challenging for GEA methods, such as when populations exhibit strong IBD (Lotterhos and Whitlock 2015) or a range expansion generates a cline in allele frequencies along an environmental gradient (Novembre and Di Rienzo 2009; Keller et al. 2009). However, in widespread species subject to a range of environments and with limited population structure, environmental and genetic distance are partially decoupled. These species can provide useful study systems for investigating spatial patterns of divergence and local adaptation.

Ideally, whole-genome sequence data would be available for characterizing the genomic basis of local adaptation as those data allow SNPs to be identified across the entire genome (e.g. (Cheng et al. 2012; Yoder et al. 2014). However, for most researchers and for most study systems, the time and money required for whole-genome sequencing, as well as the subsequent genome assembly that is needed to make sense of these genomic data, remain prohibitive. Several reduced-representation methods have been developed as alternatives to whole-genome sequencing (e.g., exome capture, ddRAD, transcriptome sequencing). For non-model species, the most popular reduced-representation approaches involve isolating random portions of the genome using restriction enzymes followed by high-throughput sequencing (see Andrews et al. 2016 a review of these methods). These methods include GSB, RAD, ddRAD, and MSG – we

will refer to these broadly as restriction site-associated DNA sequencing, following the naming used previously by Lowry et al. 2017. These approaches are widely used and are robust for a variety of population genetic analyses including estimating population structure, demographic history, and phylogeography (Catchen et al. 2013; Cavender-Bares et al. 2015; Qi et al. 2015). However, the data they produce provide only limited resolution for characterizing the genomic variants underlying local adaptation. The reasons for these limitations include: a small proportion of the genome is sampled, most markers are outside genic regions, it is difficult to compare across studies and across species, and sequenced regions are short (Tiffin and Ross-Ibarra 2014; Lowry et al. 2017). Collectively, these drawbacks limit the detecting of genomic variants that bear a signature of having contributed to local adaptation.

An alternative approach to understand the genomic basis of local adaptation is to use transcriptome sequencing, or RNA-seq data, which have been used far less frequently than other reduced representation types of data (Kozak et al. 2014; Gugger et al. 2016). RNA-seq data offer many of the advantages of restriction site-associated DNA sequencing, as data are feasible to collect in non-model organisms, and no reference genome or *a prori* sequences are needed (Wang et al. 2009). RNA-seq data are also free of many of the shortcomings of restriction site-associated DNA sequencing. Reference transcriptomes can be readily assembled through well-established pipelines, allowing the function of genes containing SNPs to be inferred; thousands of markers are generated per individual, and comparisons can be done across species (Gugger et al. 2016). Furthermore, there are analyses that can be conducted with RNA-seq data that can provide additional insight into the genomic basis of local adaptation, such as constructing co-expression networks and testing for co-expression modules that are enriched for genes that harbour signals of local adaptation. These types of analyses are not possible with data generated by restriction site-associated DNA sequencing.

Ambrosia artemiisfolia (common ragweed) is an annual, wind-pollinated plant native to North America, and widely distributed across several continents. It exhibits extensive phenotypic differentiation across both the native and invaded ranges, most notably a latitudinal cline in flowering time, with northern populations flowering earlier

and at a smaller size than southern populations (Hodgins and Rieseberg 2011; Leiblein-Wild and Tackenberg 2014; Li et al. 2014; Gorton et al. 2019). Previous population genomic work with restriction site-associated DNA sequencing in the eastern portion of the native range found weak population genetic structure (Martin et al. 2014, 2016) and identified SNP outliers associated with stress and defense, suggesting species interactions may be important to local adaptation in this species (Martin et al. 2016). However, these results were based on relatively few SNPs (2829 SNPs), therefore it is unknown whether these results are robust. Indeed, the authors themselves called for the population genetic structure and genomic basis of adaptation in North America to be further investigated with a higher density genomic dataset.

Here, we combine RNA-seq, environmental, and phenological data to examine the population genomics of local adaptation in *A. artemiisifolia*. We examined populations collected across latitude and longitude in the center of the range. In particular, we were interested in 1) assessing the utility of RNA-seq data for investigating the population genomics of local adaptation in non-model organisms, 2) characterizing patterns of genetic diversity and population structure, and 3) identifying candidate SNPs and the associated genes that may contribute to patterns of local adaptation using an environmental association analysis. We were particularly interested in using the results from the environmental association analysis to discover traits that may be important in local adaptation above and beyond those identified with traditional phenotypic approaches.

METHODS

Study system

A. artemisiifolia L. is self-incompatible (Friedman and Barrett 2008), monoecious, and wind-pollinated (Jones 1936; Essl et al. 2015) annual plant native to North America and invasive on multiple continents (Genton et al. 2005b; Chauvel et al. 2006; Xu et al. 2006). As a pioneer species, *A. artemisiifolia* is found in high disturbance, low competition habitats (Essl et al. 2015). Palynological records indicate that *Ambrosia* species were uncommon throughout eastern North America prior to the arrival of

European colonists (Grimm 2001; Williams et al. 2004). Common ragweed is now widely distributed across the continent, and population genetic estimates from contemporary samples, historical land-use data, and sedimentary pollen records suggest the expansion was most likely driven by human agricultural activity after European settlement (Grimm 2001; Martin et al. 2014, 2016).

Sample selection and tissue collection

We used seeds from 15 populations collected in Fall 2015 across a transect of 15 degrees of latitude and 7 degrees of longitude (Table 1, see Gorton et al. 2019 for additional sampling details). From these populations, we selected 4-5 maternal families per population for sequencing. We stratified the seeds from each of these families in moist silica sand and kept them in the dark at 4°C for 10 weeks to break seed dormancy (Willemsen 1975; Baskin and Baskin 1977, 1980). After stratification, we planted the seeds in 50-cell trays in BM2 germination mix (Berger, Quebec), and germinated them in a growth chamber under a 14-hour day and a day/night temperature of 22/20°C. We planted the seeds in a completely randomized design and included two to three seeds per cell. One week after planting, we randomly culled the seedlings to one germinant per cell.

We grew plants for 8 weeks, during which time we watered regularly and fertilized them once. After 8 weeks, each individual had 8-12 leaves. We harvested the top four leaves from each plant, including the apical meristem, and immediately flash froze the plant tissue in liquid nitrogen to prevent RNA degradation. Tissue was collected on a single day from 3 to 4 pm, and samples were stored at -80°C immediately after tissue collection.

RNA extraction, mRNA libraries and mRNA Illumina sequencing

We extracted total RNA using the RNeasy Plant Mini Kit (Qiagen). Prior to beginning any extraction, all equipment and counters were wiped down with 70% ethanol and RNase AWAY to eliminate RNase contamination. We extracted total RNA from the tissue samples (1 plant per family × 5 families per population × 15 populations = 75 individual plant samples) in a random order over a series of batches. Each tissue sample

was removed from the -80°C freezer, flash frozen in liquid nitrogen, and then immediately pulverized to a fine powder with a mortar and pestle previously cooled with liquid nitrogen. We then transferred a small scoop of the pulverized tissue (~ 50 mg) from each sample to a tube to begin the extraction. We eluted total RNA in 30 uL of RNase free water.

We submitted 74 RNA samples for sequencing. Sample quantification, library preparation, and Illumina sequencing were conducted by the University of Minnesota Genomics Center. Prior to library creation, each sample was quantified with Ribogreen to ensure each sample had a minimum concentration of 500 ng of total RNA. A stranded TruSeq RNA library was created for each individual plant sample. The 74 mRNA libraries were pooled and sequenced over three lanes on Illumina HiSeq 2500 with 125-bp paired-end reads (~220 million reads per lane).

Transcriptome assembly

We first constructed a *de novo* transcriptome to serve as a reference for the alignment and variant calling of the individual plant samples. After filtering out low quality reads and sequencing adapters with Trimmomatic (Bolger et al. 2014), a total of 11.8 million read pairs from a single individual (WAG1, Oklahoma) were used in the transcriptome assembly with Trinity (Haas et al. 2013). Contigs that were > 95% identical were combined with CD-HIT (Fu et al. 2012). The final transcriptome used for alignments consisted of 93,856 contigs with a median length of 719 bp and an N50 of 1484 bp. BUSCO v3 (Simão et al. 2015) was used to assess the completeness of our transcriptome assembly by searching for the presence of near-universal single-copy orthologs.

To evaluate whether the information contained in a single individual is sufficient to allow aligning reads from multiple populations, we assembled a second transcriptome using three individuals from different parts of our sampling distribution (BBM13, Minnesota; SU8, Missouri and LA429, Louisiana). Although the transcriptome assembled from three individuals included more contigs (175,144) than the single individual transcriptome, alignment proportions were highly similar with both assemblies

(96 – 98%, see next section for details). Furthermore, the BUSCO analysis indicated the three-individual assembly had lower completeness (85% complete BUSCOs for single individual assembly vs. 71% for the three-individual assembly), therefore we focused further processing and analysis on the single individual transcriptome assembly.

Individual sequence processing, alignments and variant calling

For each individual, low quality reads and sequencing adapters were removed with Trimmomatic (Bolger et al. 2014), and the surviving reads were aligned to the reference transcriptome with BWA-MEM (Li and Durbin 2010). Picard tools (https://broadinstitute.github.io/picard/) was used to add read group information and remove duplicated reads. To focus our analysis on high quality regions of the transcriptome (i.e. long contigs with high coverage) we removed reads aligning to contigs < 500 bp in length or contigs having median coverage less than 5 in a given individual (i.e. if one individual has more than 5 reads per contig, it is retained). These additional filtering steps left 42,348 contigs for variant calling and furthermore analysis.

We called variants for each individual with Freebayes (Garrison and Marth 2016), using only reads with mapping quality over 30. VCFtools (Danecek et al. 2011) was used to filter the resulting VCF-file with following requirements: site quality over 30, genotype quality over 20, minimum coverage 6·, minor allele frequency higher than 0.05, and missing data in < 20% of individuals. Indels and sites with more than two alleles were further removed.

Genetic diversity and population structure

We examined genetic diversity across all samples to obtain species-wide estimates as well as estimates within and among populations. We estimated three summary statistics with ANGSD (Korneliussen et al. 2014): nucleotide diversity π, Tajima's D and F_{ST}. Inbreeding coefficients (F_{IS}) for each individual were then estimated with PCAngsd (Meisner and Albrechtsen 2018) using a method from Vieira et al. (2013). PCAngsd also was used to assess genetic relatedness between pairs of the sampled individuals by conducting a principal component analysis (PCA) and estimating

admixture proportions. Any population structure was accounted for prior to estimating F_{IS} values.

Phenotypic and environmental data collection

All maternal sibgroups representing each sequenced population were included in a common garden experiment conducted in 2016 on the University of Minnesota campus in Minneapolis, MN (see Gorton et al. 2019 for details). In this experiment, we collected data on three different phenology traits each week. The transition to reproduction was scored as the date of the first appearance of a reproductive bud at the apical meristem, the first open male flower was scored as the date on which the first anther opened and shed pollen, and the first female flower was scored as the date at which stigmas first appeared. Due to the high correlation among the three traits (all Pearson's r >0.97), we chose female flowering time to represent flowering phenology among the populations, hereafter referred to as "flowering time". These phenotypic data were included as predictor one of the genotype-environment association analyses described below. Since selection act on phenotypic variation, we believe including phenotypic data as a variable in GEA methods is another way to identify candidate loci involved in local adaptation.

Many climate variables, including annual precipitation, annual temperature, and growing season length are highly correlated with one another as well as with latitude or longitude. Thus, to characterize the environment of the sampled populations, we used latitude and longitude of each population as proxies for climate and other geographically variable environmental factors.

LFMM2: Genome-environment association analysis

To find potential genomic targets of local adaptation, we searched for associations between SNP allele frequencies and environmental variables with latent factor mixed models in the R package LFMM2 (Caye et al. 2019). Based on cross validation error among different ridge penalty models, four latent factors (K = 4) were chosen to account for population structure in the genotype data. We conducted two LFMMs differing in the predictor variables: 1) latitude and longitude, and 2) latitude and flowering time.

Flowering time and latitude are highly correlated (Pearson's $r = -0.95$), therefore latitude was included in the second LFMM2 as additional predictor to find SNPs not already discovered by the previous analysis. For each analysis, we identified SNPs with higher than expected correlations by transforming the *P*-values to false discovery rate based *q*-values (Storey and Tibshirani 2003). SNPs with *q*-value lower than 0.05 were considered to be candidate loci which may be involved in local adaptation.

GO enrichment and functional role of candidates SNPs

To summarize the functions putatively associated with all outlier SNPs for each LFMM model, we performed BLAST queries against the *Arabidopsis thaliana* nucleotide database (The Arabidopsis Information Resource,TAIR, Berardini et al. 2015). The biological function gene ontology (GO) terms associated with the *A. thaliana* homologs (alignment *e*-value < 1×10^{-5}) were used to find common functions among contigs housing the outlier SNPs. Note that a single *A. artemisiifolia* contig may align to multiple *A. thaliana* genes, and *A. thaliana* genes can have many associated GO terms. To facilitate interpretation of these results, we summarized GO terms showing higher than expected ($q < 0.05$, Fisher's exact test) enrichment with REVIGO (Supek et al. 2011). REVIGO removes redundancies among GO terms by grouping similar terms into 'cluster representatives'. For the REVIGO analysis we used the *A. thaliana* database, with similarity set to $c = 0.5$.

GO function summaries can be coarse, overly simplistic, and often incomplete (Thomas et al. 2012; Gaudet and Dessimoz 2017). Therefore, in addition to the GO summaries, we also searched for potential functional roles of the contigs (i.e. genes) containing outlier SNPs. Using the same BLAST results as above from TAIR, we examined the functional annotations and phenotypic description for the top *A. thaliana* homolog for each match (required alignment value to *A. artemisiifolia* contig *e*-value < 1×10^{-5}). We did this for the top 50 unique contigs containing outlier SNPs for latitude and longitude, and all outliers associated with flowering time.

RESULTS

We sequenced the transcriptome from leaves and apical meristem of 74 individuals. The mean quality score for the libraries was above Q30, with a total of 144. 8 million raw reads across all libraries. The final transcriptome (assembled from a single individual, WAG1) used for aligning reads consisted of 93,866 contigs with a median length of 719 bp and an N50 of 1484 bp. After aligning the reads for all individuals to our transcriptome and removing contigs that were < 500 bp and had median coverage < 5x, we had 42,348 contigs for variant calling and downstream analysis. The proportion of reads aligning per individual after removing duplicated reads ranged from 93.7% to 98.5%, with an average of 95.8%. After removing indels and non-biallelic sites, we had 294220 SNPs which were used for our analyses with LFMM2.

Genetic diversity estimates, population structure and inbreeding coefficients

Across all samples, the species-wide estimates were: $\theta_w = 0.036$, $\pi = 0.02$, and Tajima's $D = -1.53$. The pairwise F_{ST} and PCA results indicated limited genetic structure and low genetic divergence among the sampled population, except individuals from Louisiana, which were clearly differentiated from the other samples. All populations had similar levels of genetic diversity (Table 2), with Louisiana populations harbouring slightly lower levels genetic diversity. Tajima's D estimates were negative in all populations, indicating an excess of rare variants relative to expectations under a standard neutral model.

We found that the most likely number of genetic clusters was k = 2 (PCAngsd results), corresponding to a northern cluster and southern cluster, the latter of which consisted largely of individuals from Louisiana. Populations from more central latitudes exhibited some admixture between the two genetic groups (Figure 1). The first two PCs accounted for 4.2% and 2.2 % of the genomic variance, respectively, with PC1 capturing the separation of the Louisiana individuals and structure across the range (Figure 2). We also found a pattern of isolation-by-distance, i.e. between-population F_{ST} estimates increase with geographic distance (Mantel test: p = 0.0001, Figure 3). However, as the proportion of variance explained by the PCs was low, and the pairwise F_{ST} estimates

were also low (F_{ST} = 0.06- 0.13, Table 3). Accordingly, these data show that *A. artemisiifolia* has weak population structure across the sampled range. The pairwise F_{ST} estimates were slightly higher for comparisons with populations from Louisiana (F_{ST} = 0.08 - 0.13, Table 2). Although k =2 was the most likely number of genetic clusters, we also examined the division of individuals at k =3 and k =4. The notable genetic differentiation of individuals from Louisiana was further emphasized by these results: at k = 3, the Louisiana individuals formed a genetic cluster that was distinct from the other samples, and at k = 4, one population from Louisiana (NOLA4) was further isolated (Figure 1). The separation of this Louisiana population, NOLA4, is seen along PC2 (Figure 2).

The inbreeding coefficients of most individuals were positive (average F_{IS} across all individuals = 0.04), but these values were close to zero, indicating that there was not a large heterozygote deficiency in our sampled individuals. In addition, the proportion of the sites that deviated from HWE was 20% at $p < 0.05$ and only 12% at $p < 0.01$ (using K = 2).

LFMM2 genome-environment association results, GO enrichment categories and functional annotation

Out of 294,229 SNPs, we identified 464 and 176 SNP outliers as potential targets of selection in response to environmental variation across latitude and longitude, respectively. The outliers were distributed across 256 and 125 contigs, and the GO enrichment analysis of these sequences found 176 and 101 GO terms at a frequency higher than expected by chance (Fisher's exact test, $q < 0.05$). The biological processes of the 176 terms for latitude were summarized into multiple categories including the glucose-mediated signaling pathway, plant organ senescence, sexual reproduction, and response to salt stress (Appendix 3: Figure S1). For longitude, these categories included the dephosphorylation of RNA polymerase II C-terminal domain, seed dormancy and aging (Appendix 3: Figure S2).

We identified an additional 45 SNPs, representing 42 contigs, associated with flowering time. The GO enrichment analysis of these sequences found 23 GO terms at a

frequency higher than expected by chance (Fisher's exact test, $q < 0.05$). The biological processes of these 23 GO terms were related to microtubule depolymerization, cellular response to heat, heat acclimation, and microtubule severing (Appendix 3: Figure S3).

We identified *A. thaliana* homologs for 40 and 39 of the 100 top contigs containing candidate SNPs for latitude and longitude, respectively (Appendix 3: Table S1). We identified an additional 38 homologs for 42 contigs containing candidate SNPs associated with female flowering time (Appendix 3: Table S1). The functional roles of some of these homologs in *A. thaliana* include: stomatal opening and closing (Latitude: ATOST1; Female flowering time: ATGLR3.5), circadian rhythms (Latitude: APR9), flowering time (Latitude: AP2, ELF8, PHYC, INO8O, NF-YB3: Longitude: PHYC, EMF2, RGA1; Female flowering time: ELF8, COL4, FRS10, FCA), seed dormancy (Longitude: NCED9), defense against pathogens (Latitude: CRK15; Longitude: ATNHR2B, AAO3), and stress response to abiotic factors (Latitude: OXS2, ALDH7B4; Longitude: UBP12, CPL1; Female flowering time: DREB2A, HSFA1D, SUT2).

DISCUSSION

Population genomic approaches to study local adaptation can provide complementary information to traditional organismal-based experiments by providing both historical and genetic context. For one, a clear understanding of the underlying genetic structure can contribute to disentangling which evolutionary processes may be driving patterns of divergence. Furthermore, genomic scans can identify genes that may have been subject to selection in response environmental differences without the need to select loci or traits of interest a priori. While identifying genes is not of interest *per se* (Rockman 2012; Travisano and Shaw 2013), the benefits of genomic scans arise from the ability to then use gene function to identify phenotypes which may be important to local adaptation and which may have been previously overlooked in field experiments. While RAD-seq data have increasingly been used for these genome-environment association analyses, RNA-seq data offer the opportunity to provide richer quality data and overcome many of the limitations of other reduced representation approaches.

We used RNA-seq data to identify almost 300,000 SNPs in *A. artemisiifolia*, a number that drastically exceeds previous population genetic studies of this species with RAD-seq data (Martin et al. 2016; van Boheemen et al. 2017). Consistent with previous studies of this species in North America, we found evidence for weak genetic structure across the sampled the range. This suggests that much of the phenotypic divergence found across latitude in this species (Hodgins and Rieseberg 2011; Gorton et al. 2019) may be driven by selection in response to different environmental conditions. By contrast, we did not find the deficit of heterozygotes previously found in studies of *A. artemisiifolia*, most likely due to the higher quality data offered by RNA-seq. Our GEA analyses identified putatively adaptive loci with functions involved in flowering time, abiotic stress tolerance, seed dormancy and defense against pathogens. Some of these loci suggest phenotypes that could be valuable to explore further in order to develop a more complete understanding of how selection may drive patterns of local adaptation in common ragweed.

Patterns of genetic structure and diversity in common ragweed

Previous studies have reported weak genetic structure in the native range of *A. artemisiifolia* (Martin et al. 2014, 2016; van Boheemen et al. 2017). The lack of strong population structure is not surprising given that it is wind-pollinated and produces a prodigious amount of pollen that is reported to travel potentially as far as hundreds of kilometers (Raynor et al. 1970; Lorenzo et al. 2006; Stach et al. 2007). Our results are consistent with these findings, with the majority of the sampled populations falling into one of two genetic clusters. However, at both k = 3, and k = 4, the Louisiana populations were strongly differentiated from the rest. These Louisiana populations are likely part of a distinct southeastern genetic cluster that is separate from the central US genetic cluster from which the majority of our individuals were sampled. This is further supported by the result of Martin et al. (2016), who reported a southeastern genetic cluster, encompassing individuals from Florida, Alabama and Louisiana. However, they also found that individuals from Louisiana separated out when they increased the number of genetic clusters, supporting the inference that these individuals differ ancestrally from the rest.

We are unsure what might explain these results. One hypothesis is that these populations are isolated due to differences in the selective environment and/or a lack of gene flow. While *A. artemsiifolia* is reported to be present across the southern United States, populations are much more sparsely distributed than in the midwestern portion of the range (Kartesz 2013). Thus, these Louisiana populations may not relatively disconnected, leading to greater genetic divergence between these populations and other populations across the range.

As a pioneer species, *A. artemsiifolia* thrives in high disturbance habitats including alongside roadsides, riverbanks, and agricultural fields (Essl et al. 2015). It is hypothesized to have expanded and increased in abundance across the continent in the past three centuries in part due to the spread of agricultural practices after European settlement in eastern North America (Grimm 2001; Williams et al. 2004; Martin et al. 2014). In fact, the "Ambrosia horizon" is widely used to date the layers in sediment cores and is recognized marker of European settlement (Grimm 2001; Williams et al. 2004). Martin et al. (2014) hypothesized that after increasing in abundance in the east due to human-mediated activity, western genotypes emerged from the seed bank and spread eastwards, potentially outcompeting the introduced eastern genotypes. Our estimates of Tajima's *D* are consistent with this demographic history, as our sampled populations exist entirely within the western range identified in earlier studies. All populations had negative Tajima's *D*, indicating an excess of rare variants, which is consistent with expanding population size in the west. Alternatively, our data may be detecting a much older expansion of *A. artemisiifolia* during the middle of the Holocene, during which a high amount of Ambrosia pollen has been noted (Grimm 2001; Williams et al. 2004).

Advantages of using RNA-seq data for population genomics of local adaptation

Several papers have pointed out the merits of using ddRAD and other restriction site associated DNA sequencing methods (Andrews et al. 2016; Catchen et al. 2017; McKinney et al. 2017), while others have identified limitations of using these methods, in particular for genomic scans of local adaptation (Arnold et al. 2013; Tiffin and Ross-Ibarra 2014; Lowry et al. 2017). Restriction site associated DNA sequencing methods

have been successfully used to estimate population structure, demographic history, hybridization, and migration (Cavender-Bares et al. 2015; Combosch and Vollmer 2015; Qi et al. 2015) as well as identifying quantitative trait loci (Weber et al. 2013; Lowry et al. 2014). However, there are issues with using these types of data for genome scans of adaptation. The primary issue is related to coverage: restriction site associated DNA sequencing samples only a small portion of the genome, thereby missing many potential candidate SNPs which may be involved in local adaptation (Lowry et al. 2017). This is further exacerbated for organisms with large genome sizes and/or low linkage disequilibrium (LD). The problem of sufficient coverage can be ameliorated by developing a large number of markers (i.e. RAD-tags) relative to the scale of LD, as was done in a recent genome scan on Tasmanian devils (Epstein et al. 2016). However, for many non-model organisms, the extent of LD across the genome is unknown, therefore it is difficult to assess whether the marker density is sufficient. In addition, it can be difficult to identify the genes containing the candidate SNPs with the short reads used to generate the RAD-tags. These issues, among others, suggest that restriction site associated DNA sequencing is not the optimal method for genomic scans of local adaptation, especially in non-model organisms.

RNA-seq data provides a rich opportunity to overcome some of the limitations of ddRAD for genome scans of local adaptation. The sheer quantity of data generated by RNA-seq is much higher, providing thousands of SNPs per individual genome-wide. As an example, we identified almost 300,000 SNPs in comparison to the 3000 to 10,000 SNPs identified with ddRAD in other studies on this species (Martin et al. 2014; van Boheemen et al. 2017). Although RNA-seq is itself a reduced representation method because it only samples the coding regions, it still captures many more genes than most reduced representation methods. RNA-seq data also allows one to compare the same genes across studies and across species. Nevertheless, there are some drawbacks of using RNA-seq data. Transcriptome sequencing cannot be used to investigate the importance of regulatory regions, and gene expression is both environmentally and developmentally specific. However, genic regions, or at least regions in close proximity to genic regions, are likely to be the location of much of the functional changes involved in adaptation

(Stinchcombe and Hoekstra 2008). Furthermore, there are many genes that are expressed at all times, and plants can be grown under controlled conditions to mitigate some of the environmental variation.

Candidate loci putatively involved in local adaptation in common ragweed

Flowering time is has been inferred to be important to local adaptation in multiple plant species (Jonas and Geber 1999; Fournier-Level et al. 2011; Colautti and Barrett 2013). Furthermore, latitudinal and longitudinal clines in flowering time have been previously documented in *A. artemisiifolia* (Hodgins and Rieseberg 2011; Leiblein-Wild and Tackenberg 2014; Gorton et al. 2019), and there is evidence for adaptive divergence in flowering time among populations of common ragweed (Gorton et al. 2018, 2019). Given this, it is not surprising that many of the genetic differences we identified as *A. thaliana* homologs are in genes involved in flower development and timing. Many of these were detected in the analysis with latitude as a predictor, which is consistent with the GO term enrichment of both 'sexual reproduction' and 'plant organ senescence'. While it is established that there are many flowering time genes in plants (e.g. Koornneef et al. 1998; Michaels et al. 2004; Michaels and Amasino 2007), GEA methods can refine which of those genes show evidence of importance to local adaptation. The flowering time genes we identified include PHYC and FRS10. Both of these genes are important mediators of red and far-red light responses and have important roles in flowering time regulation. PHYC has been strongly linked to altitudinal, latitudinal and longitudinal clines in flowering in other plant species (Stinchcombe et al. 2004; Samis et al. 2008; Salojärvi et al. 2017). In addition, previous genome scans have found that the spatial patterns of polymorphism in PHYC are associated with temperature and precipitation in *Arabidopsis* (Méndez-Vigo et al. 2011) and silver birch, *Betula pendula* (Salojärvi et al. 2017), providing further support for its role in climate adaptation. Salojarvi et al 2017 also identified FRS10 as a candidate gene correlated with adaptation to the environment in silver birch. Our results lend additional evidence that PHYC and FRS10 are key genes for local adaptation to latitudinal and longitudinal gradients in plants. Salojarvi et al. 2017 hypothesized that PHYC may be involved in the photoperiodic control of

inflorescence initiation in the autumn in birch. Since *A. artemisiifolia* also reproduces in late summer and early autumn, PHYC may serve a similar function in this species.

Previous work on local adaptation and range expansion in *A. artemisiifolia* has focused on selection in response to temperature (Chapman et al. 2014; Li et al. 2014), in part because this species' distribution in North America lies across a latitudinal gradient in annual temperature. Our GEA analysis of flowering time provides some further support that temperature is an important mediator of selection in this species as we found genes involved in both cellular response to heat and heat acclimation, as described by the GO term summary (Figure 6). However, these results should be interpreted with caution sine GO categories can be overly simplistic and incomplete (Thomas et al. 2012; Gaudet and Dessimoz 2017). There is also a gradient in precipitation spanning both latitude and longitude that corresponds with the edges of the distribution where *A. artemiisifolia* is most common (Kartesz 2013). This suggests precipitation may be important to both local adaptation and range limits in this species. Furthermore, previous research comparing native and introduced populations of *A. artemisiifolia* found evidence for divergence in response to drought (Hodgins and Rieseberg 2011), and Gorton et al. 2019 found the phenological responses and fitness effects of altered rainfall depended on seed source or historical climate of *A. artemisiifolia* populations. The detection of ATOST1 and ATGLR3.5 as candidate genes, both of which are involved in stomatal opening and closing (Mustilli et al. 2002; Yoshida et al. 2014), provides further support for the importance of response to water to local adaptation in this species. In addition, two recent genome scans in *Medicago truncatula* and *A. thaliana* found candidate genes involved in local adaptation that served functional roles in water stress tolerance and stomatal closure (Fournier-Level et al. 2011; Yoder et al. 2014). Collectively, these results suggest that response to water stress may be important to local adaptation in *A. artemisiifolia* and other annual plant species and indicate that variation in stomatal counts should be investigated further in natural populations.

Previous research has shown that salt may be an important driver of local adaptation to roadside habitats in *A. artemisiifolia* at relatively small spatial scales (20-30 km, DiTommaso 2004). Some of the outliers we identified as candidate genes involved in

local adaptation have functions related to tolerance to salt stress (ALDH7B4: Kotchoni et al. 2006; DREB2A: Sakuma et al. 2006), suggesting salinity may also be important to patterns of adaptation at broader spatial scales. This is not altogether surprising, given that populations in the northern portion of the range are more likely to be exposed to high salt concentrations than those from southern locations, as road salt is often applied in response to winter conditions. In addition to tolerance to salt, we also identified loci involved in seed dormancy, in particularly NCED9 (Lefebvre et al. 2006). Seed dormancy, germination timing, and germination cues are all subject to strong selection (Donohue et al. 2010; Walck et al. 2011), and dormancy expected to evolve in temperate climates where growing seasons are shorter (Rubio de Casas et al. 2017). Common ragweed is known to have strong dormancy, with seeds surviving in the seedbank up to 40 years (Toole and Brown 1946). Yet population variation in dormancy in common ragweed remains unexplored. The variable growing season lengths across the range of *A. artemisiifolia* suggest that seed dormancy could indeed be a key trait for local adaptation. Lastly, while there has been some work investigating the effects of herbivores on ragweed populations from the native and invaded range (Genton et al. 2005a; MacDonald and Kotanen 2010b,a; Gard et al. 2013), the implications of biotic interactions, or more specifically, pathogens, for local adaptation have been largely overlooked in favour of abiotic variables. Defense against pathogens has been found in genome scans of local adaptation in other plant species (Fournier-Level et al. 2011; Hancock et al. 2011; Yoder et al. 2014) and our results suggest that defense may be also be important in *A. artemisiifolia*.

Conclusions

While evolutionary ecologists may know which traits to measure to quantify fitness, we may not always know which traits are the ones selection acts upon to shape fitness. In the era of genome sequencing, instead of simply identifying genes, we think that population genomic analyses like GEA methods could be used as complementary approaches to organismal based experiments. These methods can help identify additional traits and associated environmental variables that might be driving patterns of adaptive

differentiation. For example, in common ragweed, we identified candidate SNPs involved in seed dormancy and stomatal opening and closing. Neither seed dormancy nor stomatal density have been extensively studied in common ragweed, but there is reason to expect these traits may be under selection given the biology of the species. Of course, the importance of the traits and environmental variables identified with such population genomic methods should be further investigated experimentally with common gardens and reciprocal transplants.

Chapter 4

The spatial scale of adaptation in *A. artemisiifolia* across gradients in climate

ABSTRACT

Environmental differences at both broad and fine spatial scales can influence local adaptation in plant populations. However, few experiments have sampled both geographically near and distant populations to examine the spatial scale of adaptation across a species' range and across life history stages. I conducted a multi-site reciprocal transplant experiment across latitude to investigate adaptation to broad-scale and regional-scale environmental variation in common ragweed, *Ambrosia artemisiifolia*. The experiment included 26 populations collected across gradients of temperature and precipitation, and four transplant sites in Minnesota, Iowa, Illinois and Missouri. Overall, my results suggest that the spatial scale of adaptation varies across life history stages. I found evidence for broad-scale divergence across latitude in flowering time, most likely in response to photoperiod, and evidence for regional differentiation in early survivorship. In addition, I found evidence that individuals from Minnesota may be maladapted to their native latitude, and instead, individuals from more central latitudes produced more fruits at the Minnesota site. I conclude with a discussion about the potential implications of these results for range shifts and plant responses under a changing climate.

INTRODUCTION

Decades of common garden and reciprocal transplant experiments have shown that local adaptation is widespread in plants (Turesson 1922; Clausen et al. 1940; Hereford 2009). Furthermore, it is well established that environmental differences at both broad (*e.g.*, latitudinal clines in temperature, (Agren and Schemske 2012) and fine scales (*e.g.*, soil composition, (Antonovics and Bradshaw 1970) can influence the ecological and evolutionary trajectory of plant populations, leading to changes in the phenotypic and genetic composition of populations across a species' range (Linhart and Grant 1996; Leimu and Fischer 2008). In this chapter, I combine sampling across three spatial scales with a multi-site reciprocal transplant experiment to examine the relationships between adaptive divergence, geography and environmental distance across life history stages.

The spatial scale over which adaptation occurs is expected to reflect, in part, the scale of environmental heterogeneity. However, the relevant spatial scale of environmental heterogeneity for local adaptation varies among species. For example, Brachi et al. (2013) found that selection shaping adaptive differentiation among populations of *Arabidopsis thaliana* appeared to be stronger among populations within regions than among regions, potentially due to soil factors and interspecific competition. By contrast, reciprocal transplant experiments on two perennial species *Hypochoeris radicata* and *Inula hirta* have detected local adaptation only at larger spatial scales (between regions), but not among populations within regions (Becker et al. 2006; Raabová et al. 2011). It was possible to detect the relevant scale of adaptation in these studies because the populations included in these transplant experiments were collected at more than one spatial scale, allowing both broad and regional environmental differences to be sampled. Despite the recognition that patterns of local adaptation may vary in response to the grain of environmental variation, few transplant experiments have simultaneously sampled both geographically near and distant populations to examine the spatial scale of adaptation (Fenster and Galloway 2000; Geber and Eckhart 2005; Becker et al. 2006; Brachi et al. 2013).

The degree of local adaptation depends on a balance between selection in response to environment heterogeneity and gene flow (Linhart and Grant 1996). At small

spatial scales, gene flow is expected to constrain adaptation by preventing a response to selection (Slatkin 1987; Lenormand 2002). As the distance between populations increases, environmental differences are expected to increase and gene flow is expected to decrease. Accordingly, the extent of adaptive differentiation may be expected to increase with geographical distance between populations (Montalvo and Ellstrand 2000; Joshi et al. 2001), creating a negative association between geographic distance and the degree of local adaptation. However, this pattern may be disrupted at spatial scales as small as a few meters if selection is sufficiently strong to swamp the homogenizing effects of gene flow, resulting in fine-scale differentiation (e.g. Antonovics and Bradshaw 1970). Thus, the spatial scale at which adaptation occurs can provide insight into the relative strength of the evolutionary processes underlying intraspecific differentiation.

Ambrosia artemisiifolia (common ragweed) is an appropriate system to investigate the spatial scale of adaptation because it is widely distributed across North America and experiences both broad and fine-scale environmental variation. Previous research has established latitudinal clines in flowering time in both native and introduced ranges, with plants from more northern latitudes flowering earlier (Hodgins and Rieseberg 2011; Leiblein-Wild and Tackenberg 2014; Li et al. 2014; Stinson et al. 2016; Gorton et al. 2019). These results suggest that there may be local adaptation to climate and photoperiod in *A. artemisiifolia* at geographic scales of hundreds of kilometers. However, it is unknown whether this phenotypic cline in flowering time has fitness consequences in the field as these earlier experiments were either conducted in a single common garden or did not quantify fitness. In addition, there is evidence of local adaptation to urban and rural environments between populations separated by 30-40 km (Gorton et al. 2018), suggesting that the geographic scale of gene flow and selection may occur at smaller spatial scales. As such, there is evidence that the spatial scale of adaptation may vary in *A. artemisiifolia*, but it has yet to be quantified via a rigorous reciprocal transplant experiment.

I conducted a multi-site reciprocal transplant experiment across latitude to investigate adaptation to broad-scale and regional-scale environmental variation in the native range of common ragweed, *Ambrosia artemisiifolia* (Asteraceae). The experiment

included populations collected across latitudinal and longitudinal gradients of temperature and precipitation, as well as paired sampling of nearby populations. With this experiment, I addressed the following questions: 1) To what extent is there local adaptation to climate and photoperiod in the native range of *A. artemisiifolia*? 2) To what extent are geographic distance and environmental distance of a source population from a transplant site predictive of fitness? And 3) To what extent does spatial variation at three different scales shape fitness and adaptation and how does this vary across life history stages?

METHODS

Study System

Common ragweed is a self-incompatible (Friedman and Barrett 2008), monoecious, wind-pollinated annual plant (Jones 1936; Essl et al. 2015) native to North America and widely distributed across multiple continents (Genton et al. 2005b; Chauvel et al. 2006; Xu et al. 2006). In North America, this species was uncommon in the eastern North America before the arrival of European colonists, after which its abundance increased dramatically in response to human agricultural activities spread (Grimm 2001; Martin et al. 2014). Today, common ragweed is most abundant in the eastern and midwestern United States and Canada, but it is found across both countries (Kartesz 2013). It is a ruderal plant that is often abundant in open, disturbed habitats such as river banks, roadsides, agricultural fields, and urban areas. *A. artemisiifoli*a is a summer annual that typically germinates in late spring (May-June) and flowers in late summer or early fall (August – Sept), depending on the latitude of origin. It is sensitive to freezing and can be killed by late spring and early fall frosts; the latter terminates the growing season (Essl et al. 2015).

The transition to reproduction is cued by photoperiod and is initiated when the length of day shortens sufficiently after the summer solstice. Staminate capitula (*i.e.* male flowering heads) are found in spike-like racemes, which produce pollen that is one of the primary causes of summer and fall allergic rhinitis (Lewis et al. 1983; Frenz 2001). Pistillate capitula (*i.e.* female flowering heads) are found in axillary clusters below the

male flowers; each individual flower develops into an achene (a small, single seeded fruit) that readily falls off the plant once ripe. These groups of achenes are hereafter referred to as 'fruits'.

Seed collections

In Spring 2016, I planted four common gardens separated by an average of 260 km (range: 218 km – 338 km), that span gradients in annual temperature, precipitation, and photoperiod. The common gardens, hereafter referred to as 'sites', were located in Minnesota (MN), Iowa (IA), Illinois (IL), and Missouri (MO). The MN site was located at the Rosemount Outreach and Research Center in Rosemount, MN (University of Minnesota). The IA site was located at the Conrad Environmental Research Area in Kellogg, IA (Grinnell College). The IL site was located at the Western Illinois Agricultural Field Station in Macomb, IL (Western Illinois University). The MO site was located at the Tyson Research Center in Eureka, MO (Washington University). The soil at all sites is a silt loam mixture.

In May 2016, I planted the four sites, starting with the MO site and ending with the MN site. The timing of planting at each site approximately matched the germination timing of natural populations at each latitude (Gorton, personal observations). Prior to planting, each site was sprayed with glyphosate and tilled to remove existing vegetation, creating an environment similar to the disturbed, low-competition habitats where *A. artemisiifolia* is commonly found. At each site, I planted seeds from all 26 populations. The seeds were planted in a randomized complete block design (25 blocks), so that each block contained seeds from each population (26 populations × 25 plants per population = 650 plants/field site). I selected seed material by pooling an equal number of seeds from 16-25 maternal plants per population (Table 1)

A. artemisiifolia has the potential for long-term dormancy, with seeds in the seed bank remaining viable for up to 40 years (Toole and Brown 1946). To minimize confusion between experimental and those that emerged from the seed bank, I used the ProPlugger (ProPlugger, NC) to remove a soil plug (5 cm x 5.5 cm) at each planting spot. The holes were then filled with Berger germination mix (Berger, Quebec). Prior to

planting, all seeds were stratified in moist silica sand in the dark at 4°C for 10 weeks to break dormancy (Willemsen 1975). At each planting spot, I planted 2-4 stratified seeds per maternal family directly into the ground (total number of seeds planted across all four sites = 13966) and watered the seeds immediately after planting. Each site had 65 rows and 10 columns of plants, planting spots were spaced 15 cm apart along each row, and the spacing between columns alternated between 30 and 80 cm. I planted the four field sites from May 15-May 30, 2016. Four weeks after planting, plants were thinned to a single seedling per planting spot. Throughout the growing season, all sites were regularly weeded to minimize competition from non-experimental plants.

Phenotypic and environmental data collection

I visited each site every four weeks (4, 8, 12, 16 weeks), and at the end of the growing season. During each week, I documented survivorship, plant height and scored each plant for overall reproductive stage, male flowering stage and female flowering stage (see Table 2 for scoring categories). Some plants may have germinated and died prior to the first census at 4 weeks, therefore germination (0,1) measures survivorship at 4 weeks. At 16 weeks, I also estimated female fitness by counting the number of fruits on 120-150 plants at each site. This subset consisted of 12-15 randomly selected plants from each of 10 populations, which span the full latitudinal range of the source populations: R16, R10 U25, LEW, PRC, ROI, SU8, CLW, LTR, LA4. It was not possible to quantify female fitness across whole plants due to their size and the narrow time window over which data collection needed to occur. Instead, on each plant, I counted the number of fruits on every fourth branch, starting with the lowest branch, and moving up towards the apical meristem. I subsampled branches in this way to account for potential variation in allocation to female reproduction across plant development (ca. 20-50 total branches/plant). I multiplied this number by four to get a whole-plant estimate of female fitness at 16 weeks.

At each site, I installed i-buttons (ibuttonLink, LLC) inside radiation shields to measure air temperature every hour over the course of the field season. I constructed the radiation shields out of white PVC pipe approximately 30 cm long with multiple holes

drilled throughout the pipe to allow airflow. The i-buttons were suspended on a string inside the PVC pipes, and the entire radiation shield was vertically suspended approximately 1 m above the ground.

Statistical Analyses

Timing of reproduction: All analyses were conducted in R (R Core Team 2013). Latitudinal clines in flowering time in the native range of common ragweed have been documented in common gardens (Hodgins and Rieseberg 2011; Gorton et al. 2019), but not under field conditions with varying photoperiods. To examine the extent to which timing of reproduction is predicted by the latitude of the each seed source, at each site, I fit general linear models (GLMs, *lme4* package, Bates et al. 2015) with overall reproductive stage for each plant (categorical response variable, see Table 2) at 12 weeks as a predictor and block as a random effect. I used a gaussian error distribution and an identity link function. I did not analyze the other two phenological traits at 12 weeks (male and female flowering stage), as these traits were highly correlated with overall reproductive stage (Pearson's correlation coefficient, r = 0.82-0.95, $p < 0.0001$ at all sites).

Conditional fitness components: As I was primarily interested in local adaptation, I focused my analyses on four components of fitness calculated separately for each site: probability of survival to 4 weeks, probability of flowering (female), probability of mature fruits by 16 weeks, and number of fruits at 16 weeks. I did not include any information on survival after 4 weeks because very few plants died after that time. I used female phenology scoring (see Table 2) to determine the probability of flowering and the probability of mature fruits. For the probability of flowering, all individuals that received a score of 1 or higher at any of the four censuses (*i.e.* female flowers had opened) were assigned the value of 1, and all others were assigned the value of 0. For probability of mature fruits, all individuals that received female phenology scores of 4 or higher at 16 weeks were assigned the value of 1 (*i.e.* successfully produced mature fruits), and all others were assigned the value of 0. I fit three generalized linear mixed models with different predictors to each of the four fitness components (see below). The GLMs were

conditional such that individuals were only included in the analyses if they successfully reached that life history stage (*e.g.* only plants that survived to 4 weeks are included in the analyses for probability of flowering). I used a binomial error distribution and logit link function for probability of survival to 4 weeks, probability of flowering, and probability of mature fruit, and I used a Poisson error distribution and log link function for number of fruits. Note that the analyses for number of fruits only included the 10 populations for which fruit data were collected. I conducted all analyses on individual data as well as on population means. For the individual data analyses, the GLMs for each fitness component included block as a random effect. In the chapter I focus on the individual-based analyses and report the population mean analyses in the Appendix.

To examine the relationship between geographic distance and each fitness component, I calculated the shortest distance in kilometers each field site and each source population using *distGeo* (*geosphere* package), resulting in a single value for each population, hereafter referred to as 'geographic distance'. I also tested for non-linearity in the relationship between each fitness component and geographic distance by running additional models that included geographic distance^2.

To examine the relationship between environmental difference and each fitness component, I calculated the difference in latitude between each field site and each source population, hereafter referred to as 'latitude distance'. Source populations that originated from latitudes north of a given field site had positive values, while those that originated from southern latitudes had negative values. Latitude is known to be a strong predictor of fitness (Gorton et al. 2019) and is highly correlated with many environmental variables, including climate variables and photoperiod. For example, among these sampled populations, latitude was highly correlated with mean annual temperature and mean annual precipitation (Pearson's correlation coefficient, $r = 0.97\text{-}99$, $p < 0.0001$). Thus, the difference in latitude serves as a proxy for climate and other geographically variable environmental variables.

To examine how fitness varies across spatial scales, I conducted a nested ANOVA with three different spatial scales: climate zone, region, population. 'Climate zone' refers to the USDA hardiness zone for the location of each population; 'region'

refers to the spatial grouping within a given climate zone; 'population' refers to the individual populations found within a given regional grouping (Table 1). It was not possible to run this analysis on the last fitness component, number of fruits for which we collected data on only 10 populations, an insufficient number of populations for robust comparisons. I also dropped four populations from this analysis (LA4, L46, LA1, DMS) because there was not appropriate replication of populations at the regional scale. When including block as a random effect in these analyses, the models did not converge, therefore I excluded it from the analyses.

For models with fixed and random effects, I determined the significance of all fixed effect predictors using type II Wald chi-square tests with the *car* package (Fox and Weisberg 2011), and tested the significance of block using a likelihood ratio test comparing nested models. For models with only fixed effects, all predictors were evaluated using a likelihood ratio test comparing nested models. All analyses were conducted in R, 3.5.2 (R Core Team 2018).

RESULTS

Timing of reproduction

At all sites, individuals that originated from northern latitudes were at a later reproductive stage by 12 weeks (latitude: $p < 0.0001$, Table 3, Figure 2). In other words, there was a latitudinal cline in reproductive stage, with northern populations flowering earlier than southern populations. Most individuals from the most southern latitudes were still at the vegetative stage by 12 weeks (Figure 2).

Summary of conditional fitness components

Mean survivorship at 4 weeks across seed sources was 86.6%, 82.5%, 57.4% and 77.5%, at MN, IA, IL, and MO sites, respectively. Of those individuals that survived to four weeks, mean flowering across seed sources was 72.6%, 77.4%, 78.8%, and 66.7%. Of those individuals that flowered, the mean percentage of individuals that set mature fruit by 16 weeks was 27.4%, 39.7%, 42.2%, and 36%. Of those individuals that

successfully produced mature fruit, the average number of fruits produced per individual at 16 weeks was 199, 874, 834, and 475.

Conditional fitness components: geographic distance

At IA and MO, individuals had similar survivorship at 4 weeks, regardless of geographic distance from field site (Table 4). At MN and IL, individuals from further away had lower survivorship (geographic distance: $p < 0.05$, Table 4, Figure 3), with survivorship dropping below 50% at IL at geographic distances greater than 700 km. Mean population survivorship at all sites was not significantly affected by geographic distance from field site (Table A2, A3; Figure A2, A3).

At all sites, individuals from populations closest to the site were more likely to flower than individuals from further away (geographic distance: $p < 0.0001$, Table 4, Figure 3). The probability of an individual flowering declined below 50% once individuals originated from locales greater than ~1000 km at MN, ~900 away at IA, and greater than ~700 km away at IL and MO (Figure 3). Although there was a significant non-linear relationship between the probability of individuals flowering and geographic distance at IA, IL and MO (geographic distance^2: $p < 0.05$, Table A1), the shape of relationship did not change considerably (Figure A1). At all sites, mean probability of flowering for each population declined with increasing geographic distance from field site (geographic distance: $p < 0.05$, Table A2, Figure A2), and there was no evidence of non-linearity in this relationship (geographic distance^2: $p > 0.05$, Table A3, Figure A3).

At all sites, the proportion of individuals that fruited was related to the source-distance from the field site (geographic distance: $p < 0.05$, Table 4). The two northern field sites displayed opposite patterns from the southern field sites: at MN and IA, individuals closest to the field site had the highest probability of fruiting, whereas at IL and MO, it was highest for individuals furthest from the field site (Figure 3). At MN, the probability of fruiting declined sharply from 80% to 0% over approximately 200 km: individuals from locales greater than ~400 km away did not set mature fruit. The relationship between geographic distance and probability of fruiting was less steep for the other field sites. At IA, the probability of an individual setting mature fruit declined

below 50% for individuals originating from locales greater than ~300 km away. At IL and MO, the probability of an individual setting mature fruit increased above 50% for individuals originating from locales greater than ~500 – 600 km away. There was a significant non-linear relationship between probability of fruiting and geographic distance at MN, IA and IL (geographic distance^2: $p < 0.05$, Table A1), but the overall patterns were consistent with the linear analyses (Figure A1). The population means displayed similar patterns between probability of fruiting and geographic distance (geographic distance: $p < 0.05$, Table A2, Figure A2), but there was notable non-linearity in this relationship at IA and MO (Figure A3), although geographic distance^2 was not significant at IA (geographic distance^2 >0.05, Table A3).

At all sites, the number of fruits varied with geographic distance (geographic distance: $p < 0.0001$, Table 4): individuals that originated from the nearest locales produced the most fruit at IA, IL and MO, whereas the opposite pattern was apparent at MN (Figure 3). At IA, there was significant non-linearity in the relationship between number of fruits and geographic distance (geographic distance^2: $p < 0.0001$, Table A1), whereby individuals from populations from intermediate distances produced the fewest fruit (Figure A1). The population means displayed similar patterns between the number of fruit and geographic distance, including the curvature at IA (Table A2, A3; Figure A2, A3). One notable difference was the non-linear relationship between geographic distance and the populations means at MN: populations from intermediate distances produced the most fruit (geographic distance^2: $p < 0.05$, Table A3, Figure A3).

Conditional fitness components: latitudinal distance

At MN and IL, individuals from more northern latitudes had higher survivorship (latitude distance: $p < 0.05$, Table 5, Figure 4). At IA and MO, the opposite pattern was apparent: individuals from more southern latitudes had higher survivorship (latitude distance: $p = 0.26$ at IA, $p = 0.07$ MO, Table 5, Figure 4).

When individuals originated from latitudes north of a given field site, the probability of flowering was between 80-100% (latitude distance: $p < 0.0001$, Table 5, Figure 4). By contrast, the probability of flowering declined with increasing latitudinal distance in the

southern direction. The probability of an individual flowering declined below 50% for individuals originating from southern latitudes greater than 7°-8° away at MN and IA, and greater than 3°-5° at IL and MO (Figure 4).

Similar to the probability of flowering, individuals from populations originating from latitudes north of a given field site had a higher probability of fruiting (latitude distance: $p < 0.0001$, Table 5, Figure 4). At MN, a latitude difference of only 1° in the southern direction was sufficient for the probability of fruiting to drop below 50%, and individuals from populations originating from southern latitudes greater than 3° away did not produce any fruit (Figure 4). At IA, individuals that originated from latitudes similar to the field site had a 50% chance of producing mature fruit, and this declined to 0% for individuals from southern populations greater than 5° away (Figure 4). At both IL and MO, individuals from populations at the same latitude of each field site had a low probability of fruiting (0-20%), and all individuals from populations latitudes south of MO did not produce fruit (Figure 4).

At IA, IL and MO, the number of fruits significantly declined as the difference in latitude of the source population increased (latitude distance: $p < 0.0001$, Table 5, Figure 4). By contrast, at MN, individuals that originated from more distant and southern latitudes produced more fruits (latitude distance: $p < 0.0001$, Table 5, Figure 4). The opposite pattern was apparent at the other three sites, whereby individuals from northern latitudes produced more fruit (Figure 4).

Conditional fitness components: spatial scale of adaptation

At all sites, the survivorship of individuals varied significantly across all three spatial scales (Table 6, Figure 5). With the exception of IL, individuals from the most southern hardiness zone had the highest probability of surviving. Individuals originating from eastern regions within the two most northern hardiness zones (zone 4,5) had an equal or higher probability of surviving than individuals from the western region (Figure 5). The exception was individuals from hardiness zone 5 at IA, whereby individuals from the western region had higher survivorship. At MN, IA, this resulted in individuals originating from the region nearest the field site having the highest probability of

surviving (MN: region 4x = 95%, IA: region 5x = 83%, Figure 5). The pattern of survivorship among regions within climate zone 6 and 7 were not consistent across sites, nor among zones (Figure 5). At MO, individuals originating from the region nearest to the field site had the lowest survivorship (MO: region 6x = 59%, Figure 5). The range of survivorship among individuals from different populations within a region was considerable: at MN, the individuals from different populations within a region had a mean probability of survivorship that ranged from 12-99%, at IA it was 36-96%, at IL it was 12-88%, and at MO it was 26-96%.

Hardiness zone was the primary spatial scale that explained the probability of flowering: individuals from the three northern hardiness zones were more likely to flower than individuals from the most southern zone (zone: $p < 0.0001$ at all sites, Table 6, Figure 6). There were no significant differences among regions in the probability of flowering at MN, IL or MO (region: $p > 0.05$, Table 6). At IA, individuals from region 6y, had a 38% lower chance of flowering compared to the other regions in that climate zone (region: $p = 0.06$, Table 6). At MO, there was evidence for variation among populations in the probability of flowering (population: $p = 0.02$, Table 6), but the deviance explained by population was considerably lower than climate zone.

At all sites, individuals from populations originating from the most northern hardiness zone had the highest probability of producing mature fruits by 16 weeks (zone: $p < 0.0001$, Table 6, Figure 7). At MN, IA and IL, individuals from populations originating from more eastern regions were more likely to successfully produce mature fruits (region: $p < 0.001$, Table 6, Figure 7), whereas at MO, only hardiness zone affected the success of fruiting (region: $p = 0.94$, Table 6, Figure 7).

The population means across all three fitness components showed little variation among spatial scales, with the exception of climate zone (Table A5).

DISCUSSION

The grain of environmental heterogeneity that shapes patterns of local adaptation varies across species. To understand the scale of local adaptation within a species, I conducted a broad scale reciprocal transplant experiment with a wide-spread, wind-pollinated,

summer annual. Overall, my results suggest that the spatial scale of adaptation varies across life history stages. Specifically, I found 1) geographic distance does not always predict adaptation, 2) broad-scale adaptive differentiation of flowering phenology, 3) regional differentiation in early survivorship, 4) and microscale variation for fruit numbers. Below, I discuss the relationship between local adaptation and geographic distance, discuss patterns of differentiation at each of three spatial scales, and speculate about the implications for range shifts and climate change.

Geographic distance, environmental distance and local adaptation

One of the general findings of this experiment is that geographic distance is not always an adequate predictor of adaptive divergence. While the analyses with geographic distance showed the expected pattern (i.e. adaptive divergence increases with increasing geographic distance), the latitudinal analyses, which used a polarized predictor, revealed a more nuanced pattern. These results suggest that geography alone can fail to capture the environmental differences that can covary with distance.

The scale of local adaptation can be investigated by comparing distant vs near populations, with geographic distance as a predictor of ecologically important traits or fitness (Galloway and Fenster 2000). This is reasonable as environmental differences are likely to increase and gene flow is likely to decrease with increasing geographic distance, thus it might be expected that the extent of adaptive differentiation would also increase with distance. However, populations separated by both small and large geographic distances display adaptive divergence across a variety of ecologically important traits including flowering time, size, and heavy metal tolerance (Antonovics and Bradshaw 1970; Jonas and Geber 1999; Agren and Schemske 2012). Furthermore, the extent of divergence in these traits is not always explained by the geographic distance separating the populations of interest.

Given the limited genetic structure in *A. artemisifolia* (Ch. 3), we know that it is unlikely that the observed patterns of phenotypic differentiation are caused only by decreased gene flow and increased drift across geographic distance. In fact, it is environmental differences, and selection in response to these environmental differences, that leads to

adaptive differentiation. My analyses with latitude (a proxy for climate, photoperiod, and other environmental variables) illustrate this important point. Individuals that were collected from the same distance away from a given field site showed very different patterns in flowering and fruiting when considering latitudinal distance. Individuals from latitudes north of a given field site generally flowered and produced mature fruit, while those from latitudes south of a given field site did not. If geographic distance was the metric that should be used in local adaptation studies, such directionality would not be important. These data demonstrate that using geographic distance alone is not sufficient to examine patterns of adaptive divergence. Instead, the use of environmental distance metrics – estimated from either publicly available data (e.g. WorldClim, Harmonized World Soil Database) or collected in the field by researchers – calculated at the spatial scale of interest should be used.

Adaptation and divergence across spatial scales in common ragweed

At broad-spatial scales, my results suggest that ragweed as adapted to varying photoperiod via differentiation in flowering time. I found a strong latitudinal cline in the timing of flowering with individuals from northern latitudes flowering earlier than individuals from southern latitudes at all four sites. More strikingly, individuals from the most southern latitudes (Arkansas, Louisiana) typically flowered too late (if at all) to successful produce mature fruit in any of the four sites. These results are consistent with single-site common garden studies of *A. artemisiifolia* (Hodgins and Rieseberg 2011; Leiblein-Wild and Tackenberg 2014; Gorton et al. 2019), as well as, work showing that flowering time in ragweed is cued by daylength (Deen et al. 1998). These results also mirror patterns of flowering time and bud set observed in other species that have been shown to be primarily cued by daylength (Lacey 1988; Li et al. 2003; Luquez et al. 2008; Keller et al. 2011). While photoperiod and climate are confounded in this experimental design, given the natural history of ragweed and the similarity of these results with other daylength sensitive species, it is reasonable to assume that photoperiod plays a significant role in broad-scale patterns of adaptation in *A. artemiisifolia*.

In contrast to the importance of broad-scale environmental differences to shaping the probability of flowering and fruiting in *A. artemisiifolia*, early survivorship varied at all three spatial scales. While we found variation across spatial scales for most fitness components, it was consistently found for early survival. In particular, I found support for adaptive divergence at a regional scale, as individuals originating from the regions nearest to the MN and IA field sites had the highest survivorship. This suggests there has been local adaptation at smaller scales for early life history stages such as survival. It is reasonable to expect that germination and early survivorship is under selection in *A. artemisiifolia* as the timing of seed germination, germination cues and dormancy are all subject to strong selection (Donohue et al. 2010; Walck et al. 2011). Furthermore, adaptation in seed traits can shape later life history stages (Burghardt et al. 2015). More broadly, these results suggest that the spatial scale of adaptation varies across life history stages.

As already described, latitudinal distance from a field site has a strong effect on phenology. However, for those plants that are able to successfully flower and set mature fruits, microenvironmental variation within a field site has a considerable effect on the number of fruits produced. The magnitude of the effect of within site environmental effects is apparent from the large deviance of block for the number of fruits at all field sites (Table 4). It is difficult to determine what may be driving these patterns. One potential explanation is the use of a controlled environment for this experiment. I regularly weeded the field sites, thereby controlling competition from non-experimental plants. However, competition in common ragweed can shape patterns of survival and fitness (MacDonald and Kotanen 2010b). Given that the biotic environment likely varies across the range, it's possible that allowing more 'natural' conditions may have resulted in more pronounced effects of seed source on fruit production, although it is unlikely to lessen the environmental variance. These results indicate that microhabitat variation or localized environmental variation could override regional or continental differences and be essential in determining fruit production in this species. Moreover, these results emphasize the importance of considering the magnitude of effect in experiments and not simply statistical significance.

Study limitations

There are a few caveats to this work. I analyzed the components of fitness separately, but fitness over an entire life span (i.e. lifetime fitness) must be assessed by integrating these components together. I plan to use aster modelling (Geyer et al. 2007; Shaw et al. 2008) to examine lifetime fitness. I also used latitudinal distance as a proxy for environmental distance. Moving forward, I plan to explicitly calculate environmental distance using precipitation and temperature data from WorldClim 2.0 (Fick and Hijmans 2017) and soil quality and land cover data from the Harmonized World Soil Database 1.2 (Fischer et al. 2008). The other issue is related to seed maturation and viability. Although individuals from mid latitudes may successfully produce fruit when planted at northern latitudes, whether or not these fruits fully mature prior to the kill frost is another question. Seeds can take 4-6 weeks to develop to maturity following pollination (Essl et al. 2015). Similarly, at the southern field sites, IL and MO, there was up to an additional 8 weeks of the growing season after the 16-week census, which would have contributed to the maturation of additional fruits on some of the southern seed sources. Ideally, fruits should be collected from all individuals and then germinated to truly assess fitness. However, this is difficult in *A. artemisiifolia*, as fruits readily fall off plants 1-2 weeks after ripening (Essl et al. 2015). Lastly, I plan to analyze the four gardens together to test for interactions between each spatial scale of interest and common garden.

Implications for climate change and range shifts

Broad-scale adaptive divergence in flowering time has important potential implications for how common ragweed will respond to climate change. At higher latitudes, average temperatures are cooler, and the length of the growing season is shorter in comparison to southern latitudes. In response, plant populations at higher latitudes have evolved to flower and set seed earlier, which results in a completion their life cycle before the first frost (Jonas and Geber 1999; Olsson and Ågren 2002). My results, as well as, others support this expected pattern in common ragweed (Hodgins and Rieseberg 2011; Leiblein-Wild and Tackenberg 2014; Gorton et al. 2019). Early flowering, however, can have negative fitness consequences, as plants will have fewer resources

available for subsequent growth and reproduction (Reekie and Bazzaz 1987; Geber 1990; Dorn and Mitchell-Olds 2006). The fitness cost of early reproduction was apparent in this experiment: individuals from northern latitudes flowered earlier at all sites, and these individuals also senesced earlier and produced fewer fruits. In real terms, the cost of early reproduction is expected to constrain the response of local populations to climate change, as plants, such as common ragweed, that are locally adapted to given photoperiod, will continue to reproduce at the same time of year, despite the extended growing season (Way and Montgomery 2015). Indeed, this maladaptive consequence of climate change might explain a curious result from this study. While individuals at IA, IL, and MO all display the expected pattern of local adaptation—as geographic distance from the field site increases, fitness declines— individuals from MN exhibited the opposite pattern, whereby individuals from the native latitude had the lowest fitness. In Minnesota, the growing season has become significantly longer over the century due to climate change, meaning the local photoperiod is now mismatched with the length of the growing season (Wuebbles et al. 2017). As a result, individuals from southern latitudes that cue on a shorter critical photoperiod capitalize on the longer growing season and produce more fruit when grown in Minnesota. The lower fitness of native Minnesota population grown in Minnesota relative to their southern counterparts may reflect a lag in the migratory response of ragweed to climate change, a response that has been predicted/found in other annual species (Wilczek et al. 2014).

While MN is not the northern geographic limit of *A. artemisiifolia* (Genton et al. 2005a; Kartesz 2013), there are interesting implications for the northward migration and spread of common ragweed. My results suggest that individuals from populations from the middle of the range may migrate northwards and thrive under climate change. Northern populations might also evolve extended flowering periods in response to the longer growing season and thereby increase fruit production. It has not been established if there is genetic variation for the cessation of flowering, which would be required in order for these populations to evolve later flowering. However recent data indicate that ragweed pollen season has extended at northern latitudes (Ziska et al. 2011), providing

some evidence that these populations are responding to the longer growing season, whether through plasticity of genetic change or both.

The extent of adaptation at different life history stages and at different spatial scale may also limit the potential for this species to migrate. For example, while southern populations might be able to capitalize on the extended growing season due to a shorter critical photoperiod, regional conditions required for germination or early survivorship might not be found in more northern climate zones. Indeed, the potential for different environmental factors to drive adaptation of different life history stages and constrain responses to climate change has been seen in other systems, both within (Lehikoinen et al. 2009) and among species (Kimball et al. 2010).

Conclusions

One of the broader implications of these findings is that plants may require multiple life stage adaptation. While it may be easier to obtain this combination of alleles in a wind-pollinated, outcrossing species such as common ragweed, in a selfing species, it could be much more difficult. Furthermore, while we may not be interested in restoring or maintaining populations of ragweed, the principles here are broadly applicable to other plant species, in particular other summer or spring annuals. When examining adaptation across spatial scales, it becomes apparent that adaptation at different life history stages may contrast and potentially limit plant responses, with implications for their ability to respond to climate change.

Figures – Chapter 1

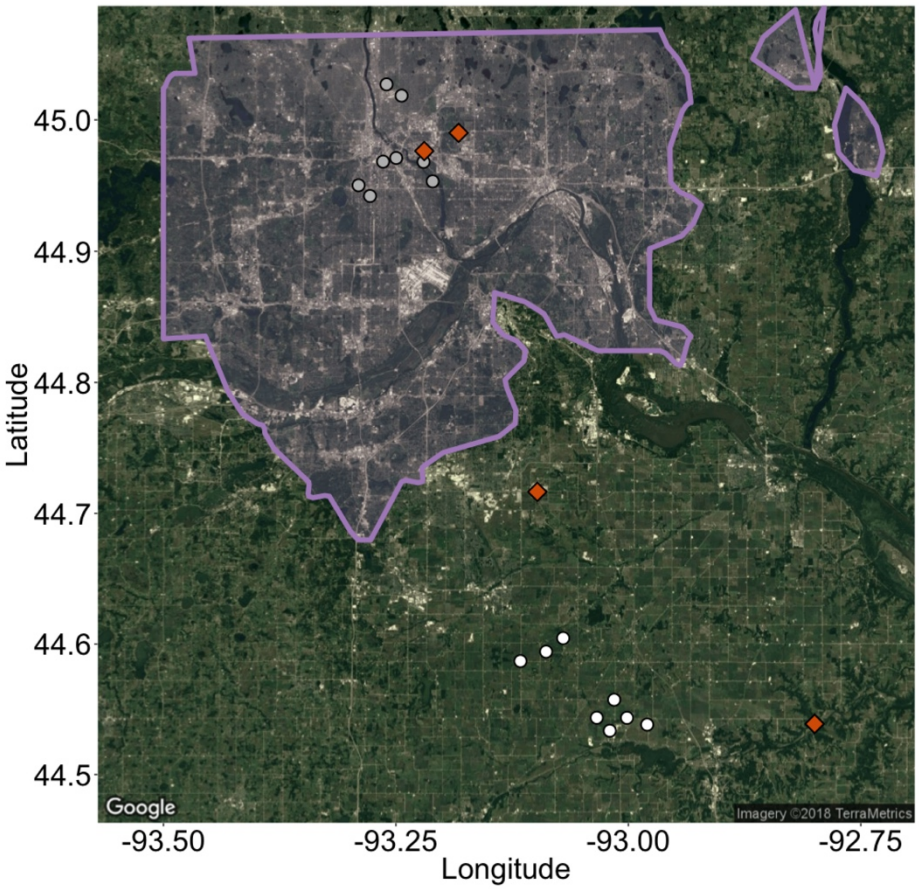

Figure 1. Map of sampling locations and field sites in the Minneapolis area. Urban sampling locations are indicated in grey, rural sampling locations are indicated in white. Common garden locations are indicated in red. The purple outline indicates urban areas defined by MODIS satellite data at 0.5 km resolution (Schneider et al. 2009).

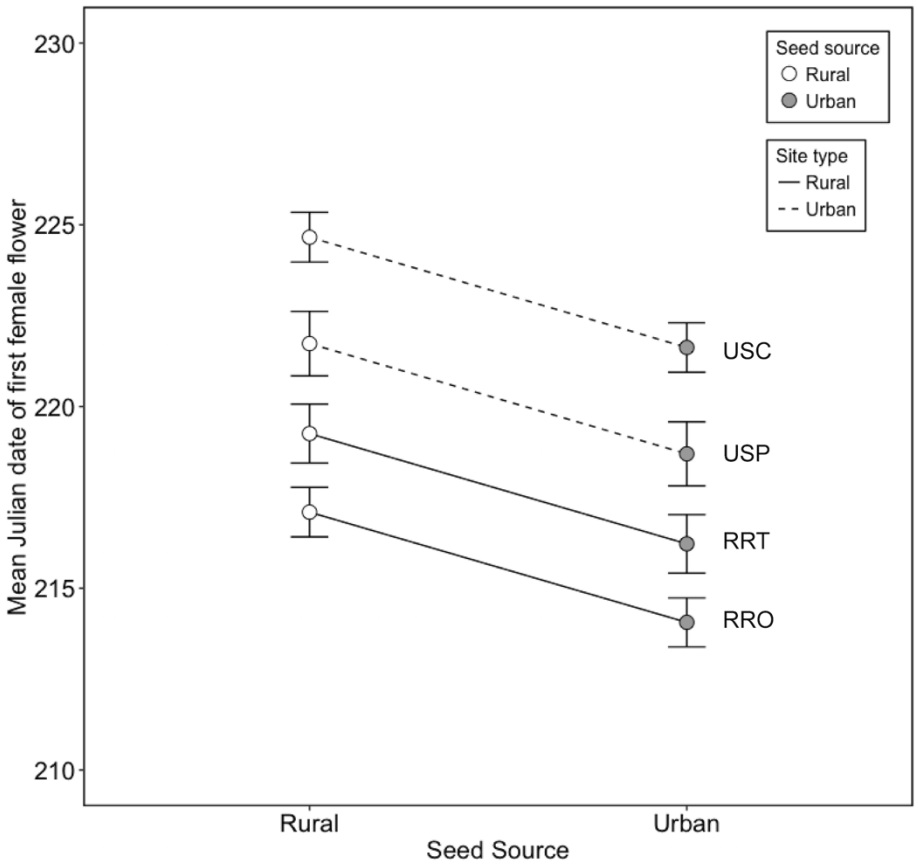

Figure 2. Mean Julian date of first open female flower for urban and rural populations. Grey = urban populations, white = rural populations**.** Dashed lines = urban field sites, solid lines = rural field sites. The bars indicate +/- standard error. There was a significant effect of regional seed source ($p = 0.0006$) and site ($p = < 0.0001$) (Table S3). Means were extracted from linear mixed effects model.

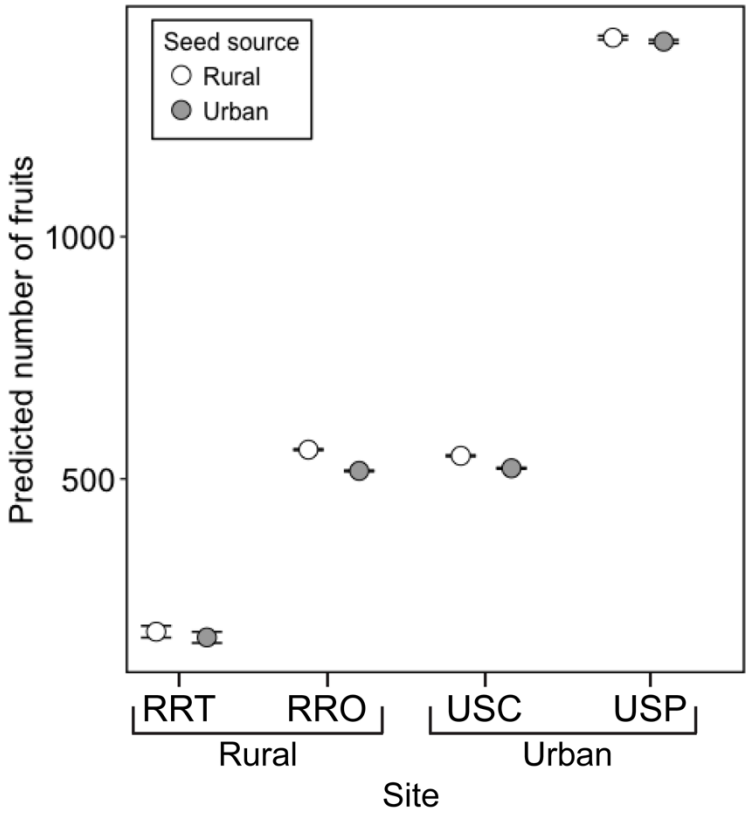

Figure 3. Mean number of predicted fruits produced by urban and rural populations at each field site. White = rural populations, grey = urban populations. The bars indicate +/- standard error. Predicted means were extracted from separate aster models for each site. There was a significant effect of source region (see Table S9). Rural field sites = RRT, RRO; urban field sites = USC, USP.

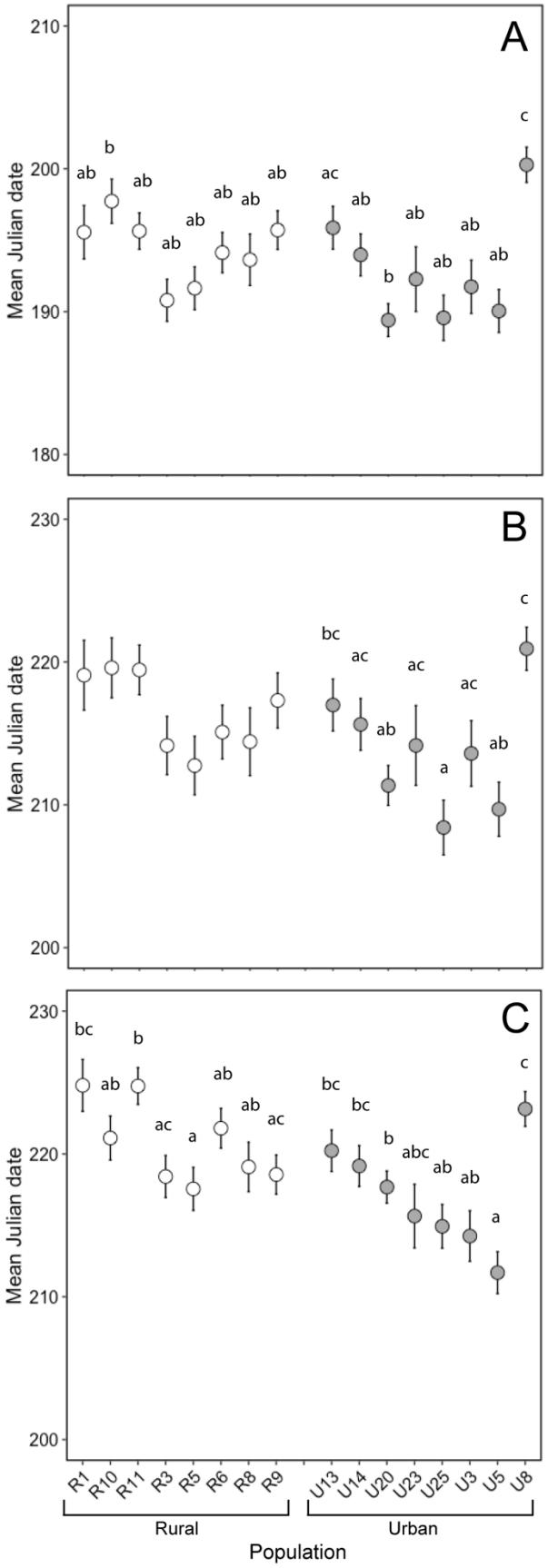
A
B
C
Mean Julian date
210
200
190
180
230
220
210
200
ab
b
ab
ab
ab
ab
ab
ab
ac
ab
b
ab
ab
ab
c
bc
ac
ab
ac
a
ac
ab
c
bc
b
ab
ac
a
ab
ab
ac
bc
bc
b
abc
ab
ab
a
c
R1
R10
R11
R3
R5
R6
R8
R9
U13
U14
U20
U23
U25
U3
U5
U8
Rural
Urban
Population

Figure 4. Mean Julian date of A) transition to reproduction, B) first open male flower, and C) first open female flower of urban populations and rural populations averaged across field sites. White = rural populations, grey = urban populations. The bars indicate +/- standard error. Letters indicate Tukey test results conducted among rural or among urban populations, where unique letters indicate significantly different groups. Urban and rural population names are listed along the x-axis in an arbitrary order based on the numerical identity of the population. There was a significant effect of population seed source on all flowering time variables (Table S6). Means were extracted from linear mixed effects models for each trait.

Figures – Chapter 2

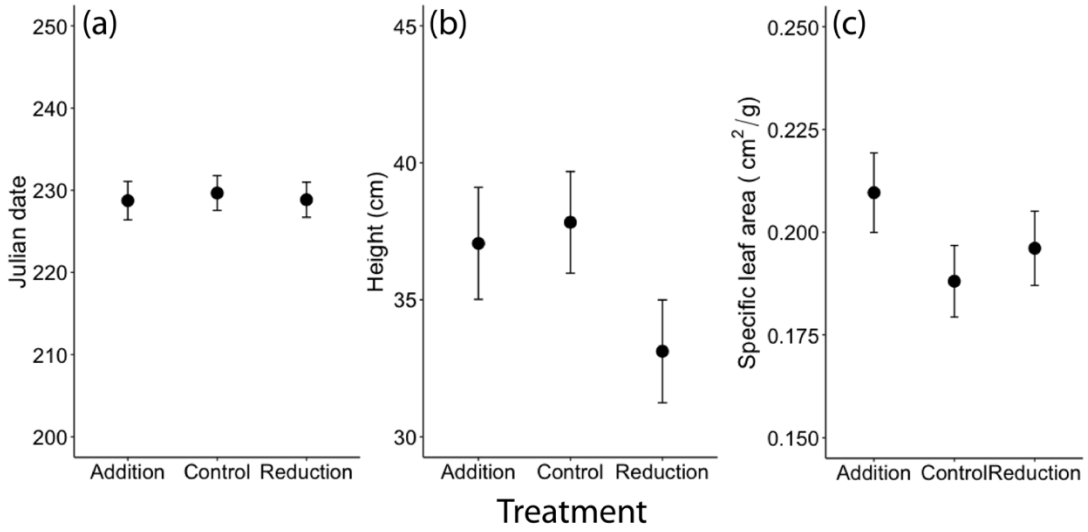

Figure 1. Effects of rainfall treatment on a) mean julian date of transition of reproduction, b) mean height at 8 weeks, and c) mean specific leaf area. Least-square means were extracted from linear models and are presented with standard error bars.

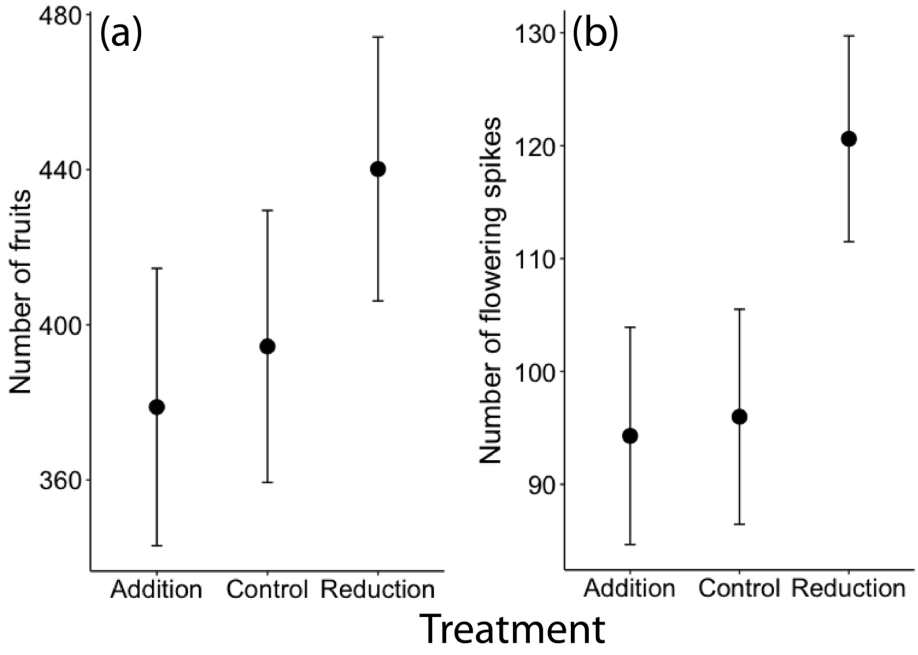

Figure 2. Effects of rainfall treatment on a) mean number of fruits (female fitness), and b) mean number of flowering spikes (male fitness). Predicted means were extracted from *aster* models and are presented with standard error bars.

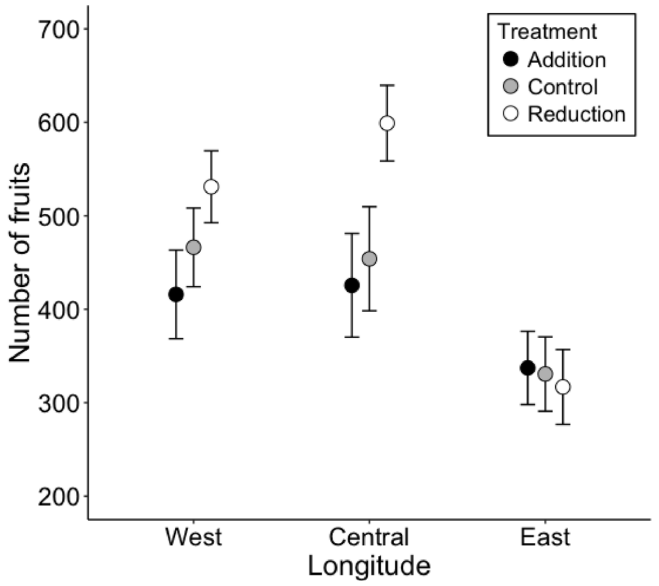

Figure 3. Effects of rainfall treatment and longitude on mean number of fruits (female fitness). Predicted means were extracted from aster models and are presented with standard error bars.

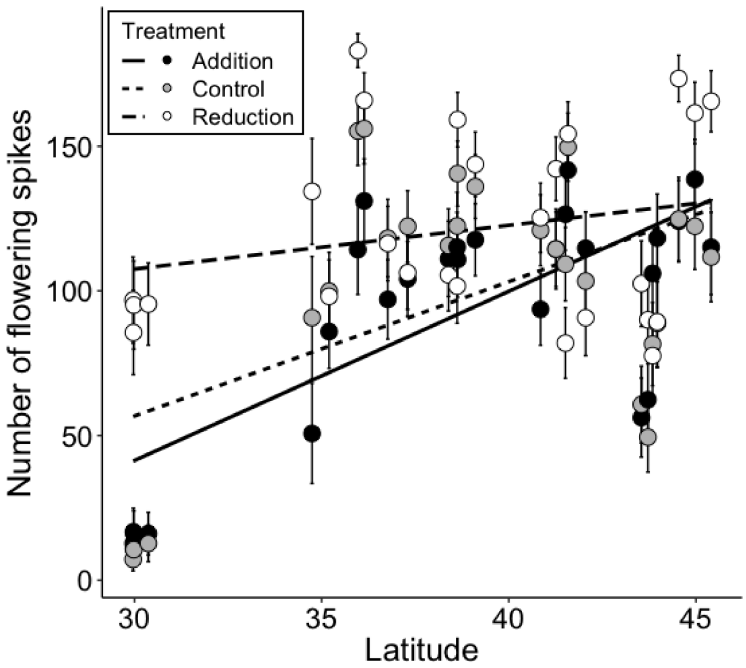

Figure 4. Effects of rainfall treatment and latitude on mean number of flowering spikes (male fitness). Predicted means were extracted from aster models and are presented with standard error bars. Linear regression lines are shown for male fitness vs latitude for each rainfall treatment.

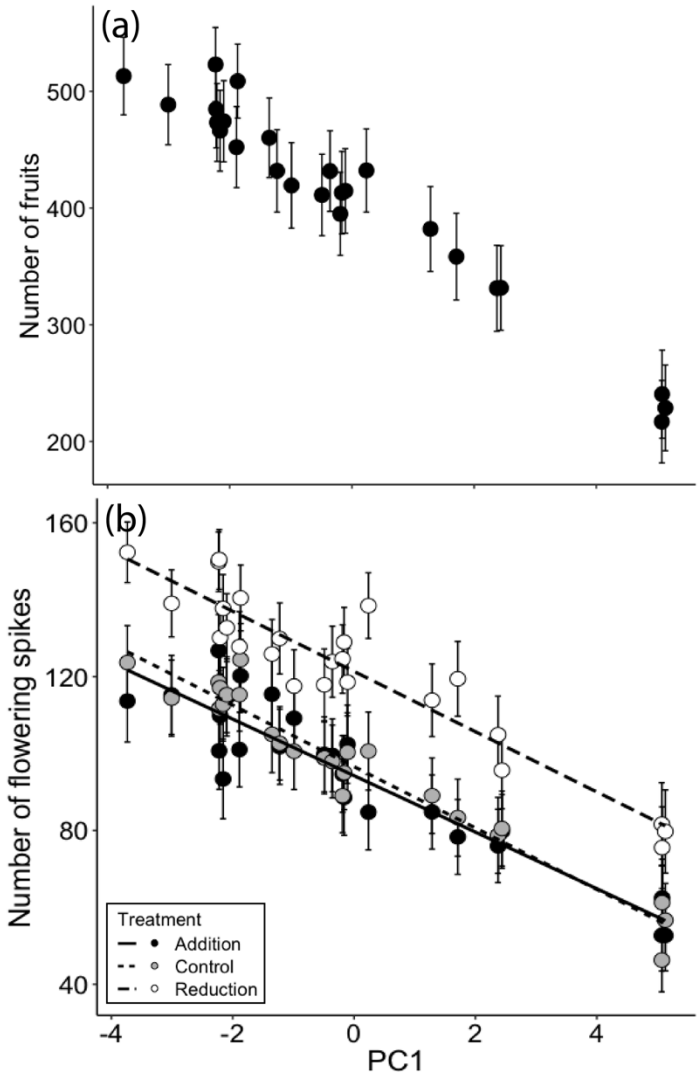

Figure 5. Effects of PC1 of BIOCLIM precipitation values on a) mean number of fruits (female fitness), and b) mean number of flowering spikes (male fitness). Linear regression lines are shown in b) for male fitness vs PC1 for each rainfall treatment. Predicted means were extracted from *aster* models and are presented with standard error bars.

Figures – Chapter 3

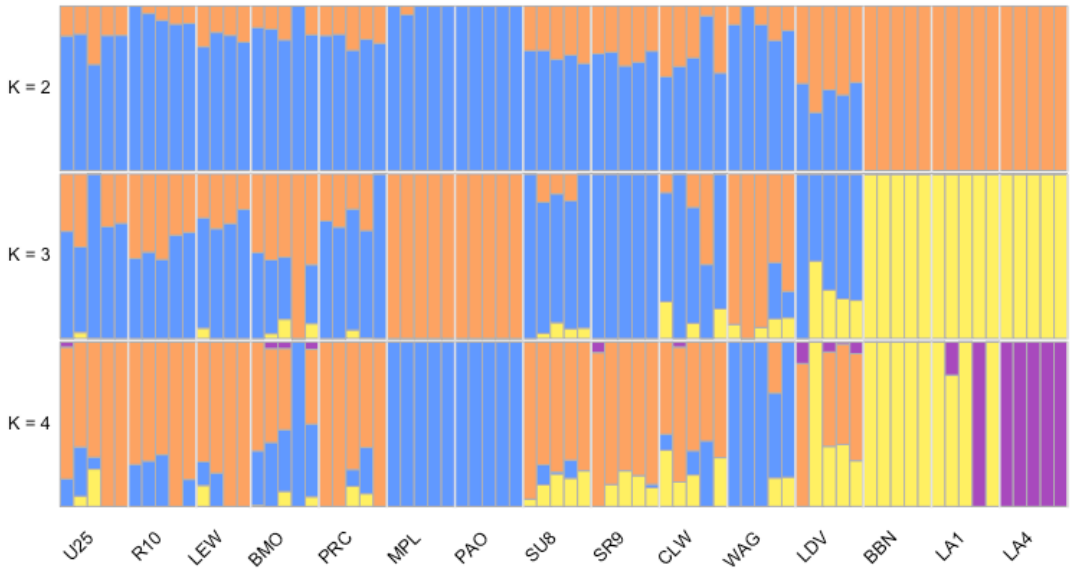

Figure 1. Estimated admixture proportions (PCAngsd) for three different numbers of ancestral populations (K). Each bar represents a single individual. Population names are shown along the horizontal axis, and are listed in latitudinal order.

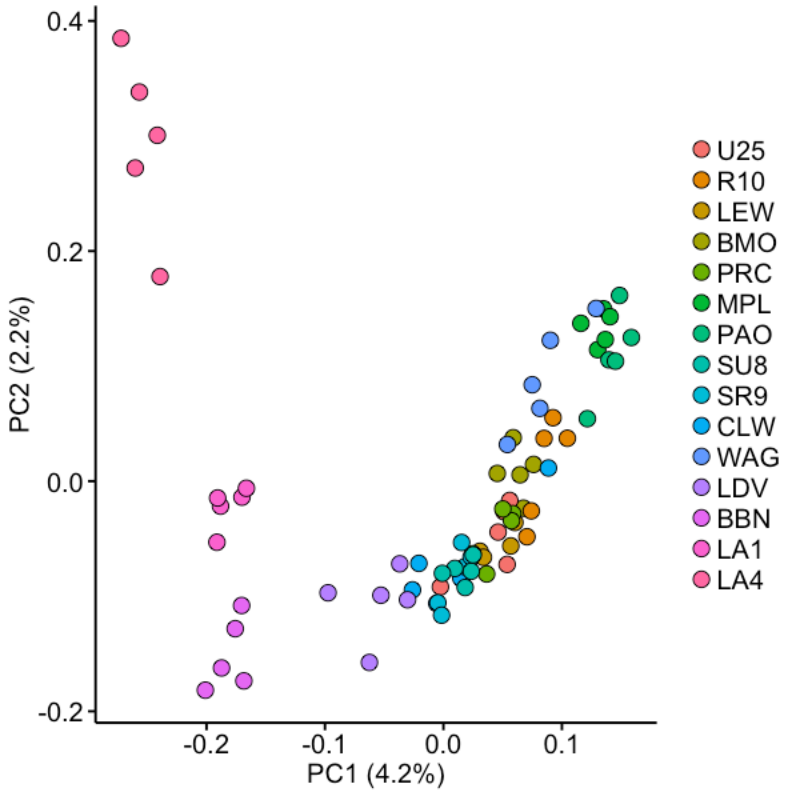

Figure 2. Genetic variation along the first two axes of a principal components analysis (PCA). Variation explained by each PC is shown in brackets. Populations are listed in latitudinal order, starting with the most northern populations. Each colour represents individuals from a different population.

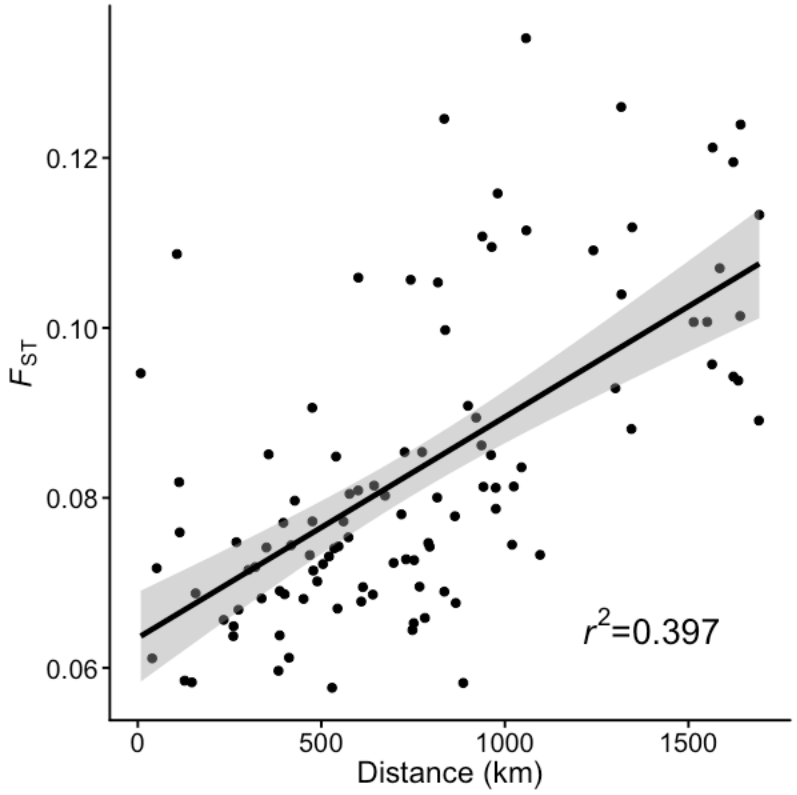

Figure 3. The linear relationship between genetic distance (F_{ST}) and geographic distance between populations.

Figures – Chapter 4

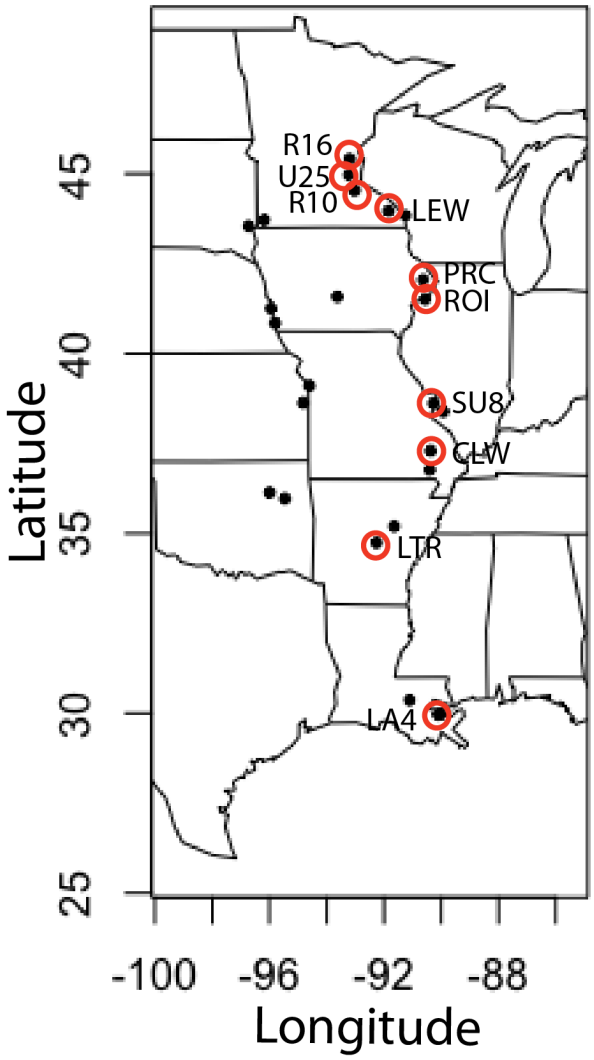

Figure 1. Location of field sites and source populations. The black dots indicate the location of the 26 source populations (see Table 1 for exact coordinates). The populations circled in red indicate those populations with fruit counts at 16 weeks. The location of each site is indicated with a blue triangle.

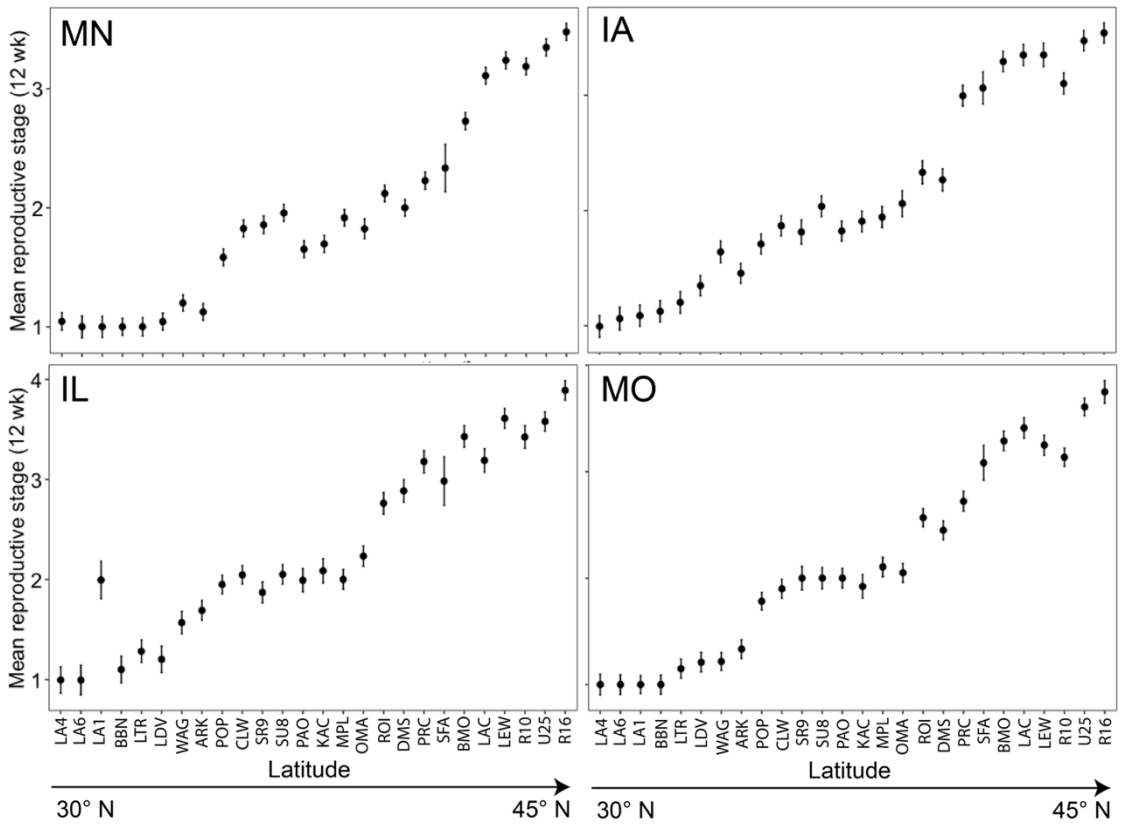

Figure 2. The effect of source latitude on overall plant reproductive stage at 12 weeks. Values above 3 indicate either male or female flowers were open. See Table 2 for a description of all potential stages. Least-squared means were extracted for each population at each site. Populations are listed in ascending order by latitude. Standard error bars are shown.

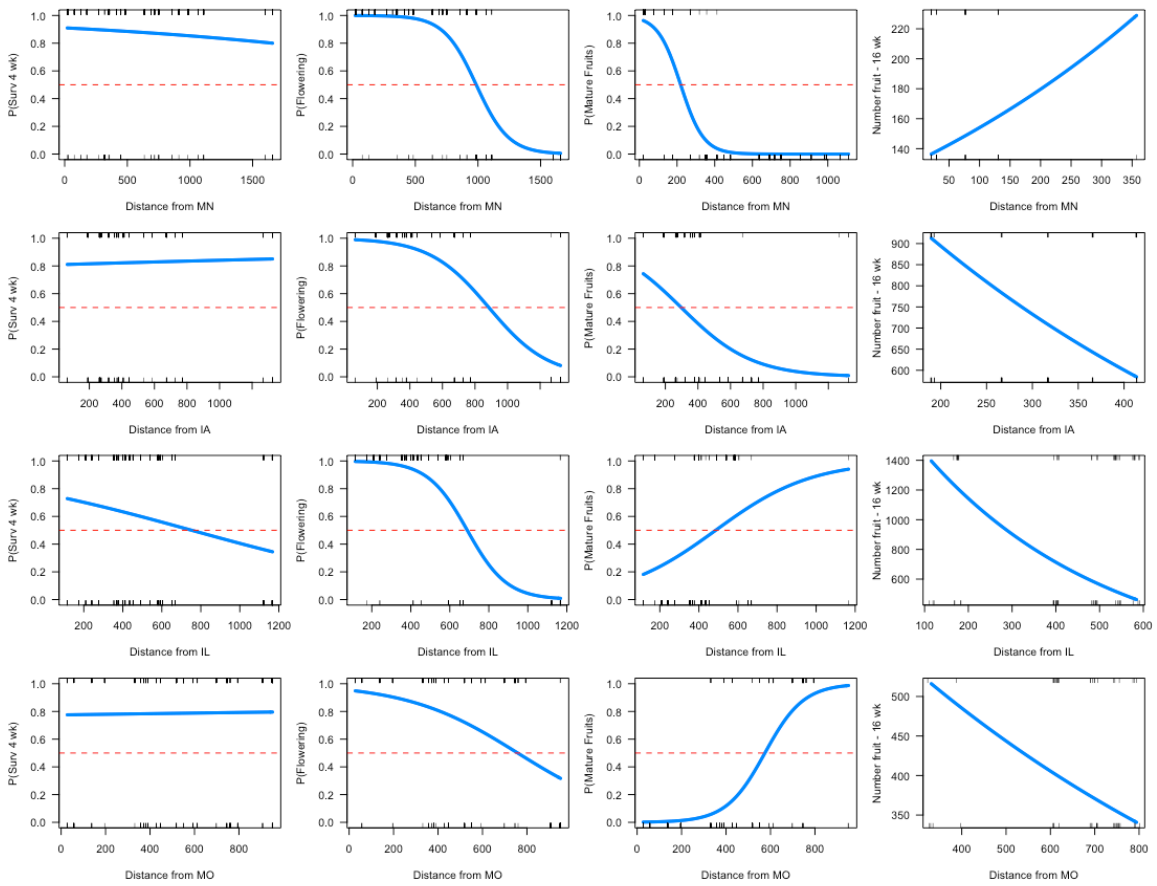

Figure 3. The relationship between fitness and geographic distance of source population from each of the four transplant sites. Generalized linear mixed models were run on individual data and included block (random) and geographic distance from field site (fixed) as predictors. The model outcomes were transformed back to the response scale. Each row is a different site (MN, IA, IL, and MO), and each column is a different fitness component (P(Surv 4 wk), P(Flowering), P(Mature Fruit), Number of Fruit at 16 weeks). The blue line represents the expected value, and red dotted line indicates where the probability is 0.5. There are no confidence intervals due to the inclusion of block as a random effect.

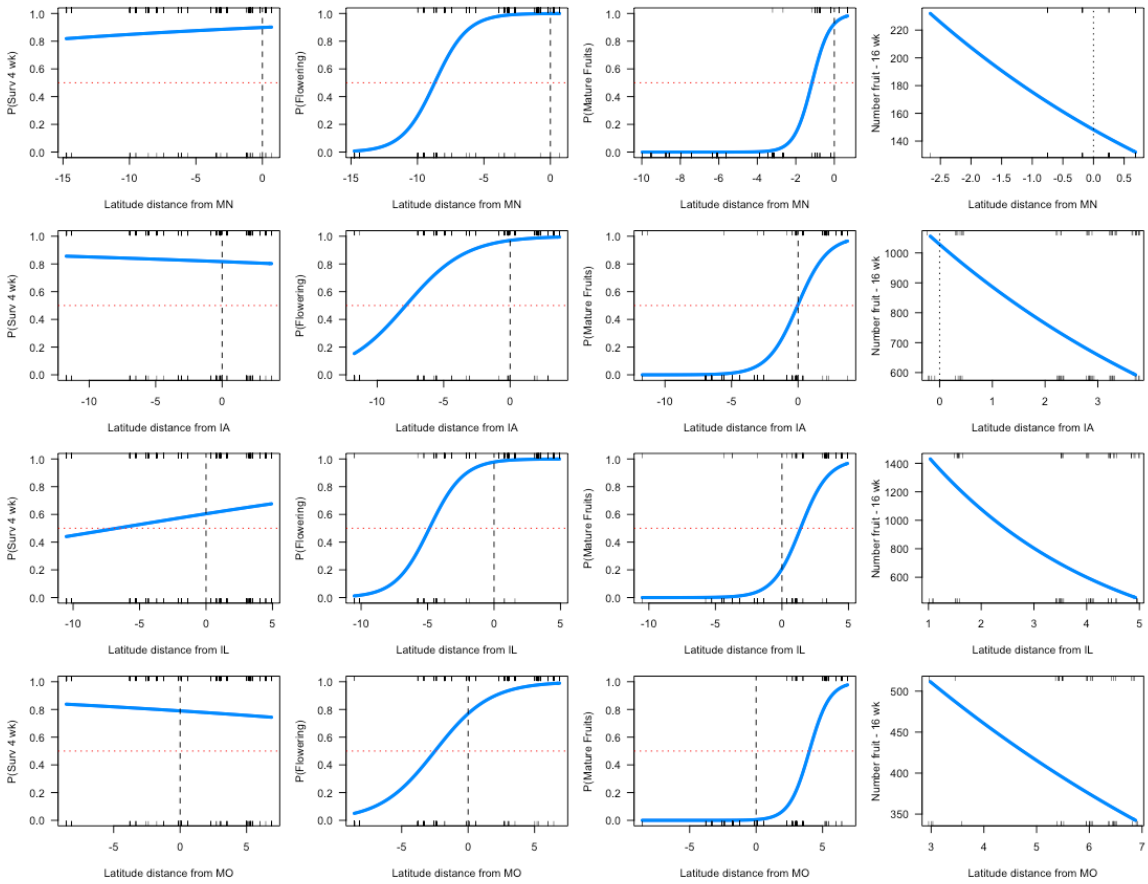

Figure 4. The relationship between fitness and latitudinal distance of source population from each of the four transplant sites. Generalized linear mixed models were run on individual data and included block (random) and latitudinal distance from site (fixed) as predictors. The model outcomes were transformed back to the response scale. Each row is a different site (MN, IA, IL, and MO), and each column is a different fitness component (P(Surv 4 wk), P(Flowering), P(Mature Fruit), Number of Fruit at 16 weeks). The blue line represents the expected value, and red horizontal dotted line indicates where the probability is 0.5, and the black horizontal dotted line represents the transplant site location. There are no confidence intervals due to the inclusion of block as a random effect.

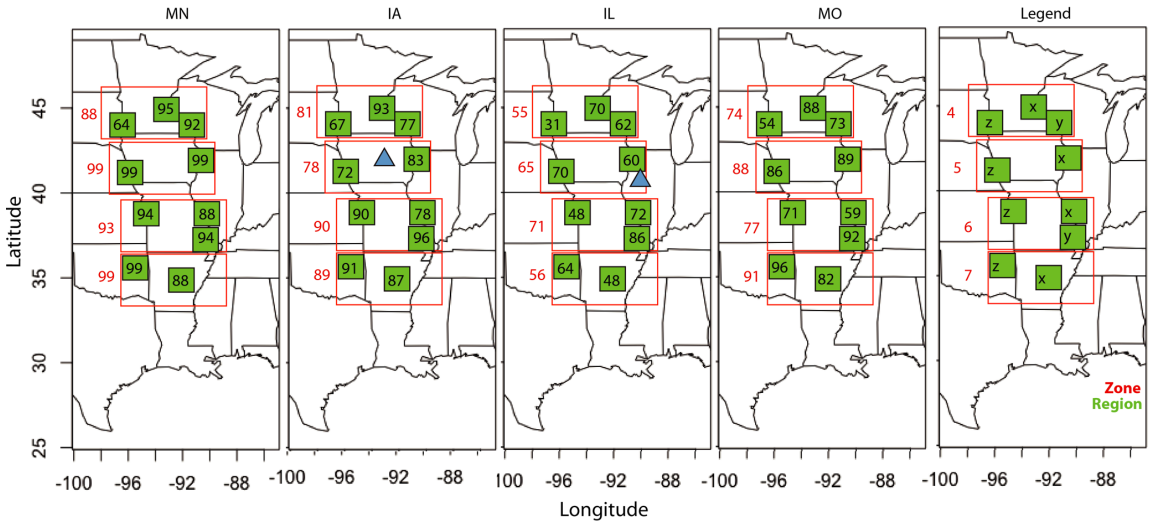

Figure 5. Mean probability of survival at four weeks across spatial scales, corresponding to the results in Table 6. Each red box and text delineate a separate hardiness zone and the associated mean probability of survival for individuals from that zone. Each green box is a regional grouping within a zone, and the number indicates the mean probability of survival for individuals from that region. The legend on the right specifies the names for each zone and region, and the same legend applies to Figure 5 and 6. Blue triangles indicate the location of each site. Note that for MN and MO, the location of the site is masked by the regional groups.

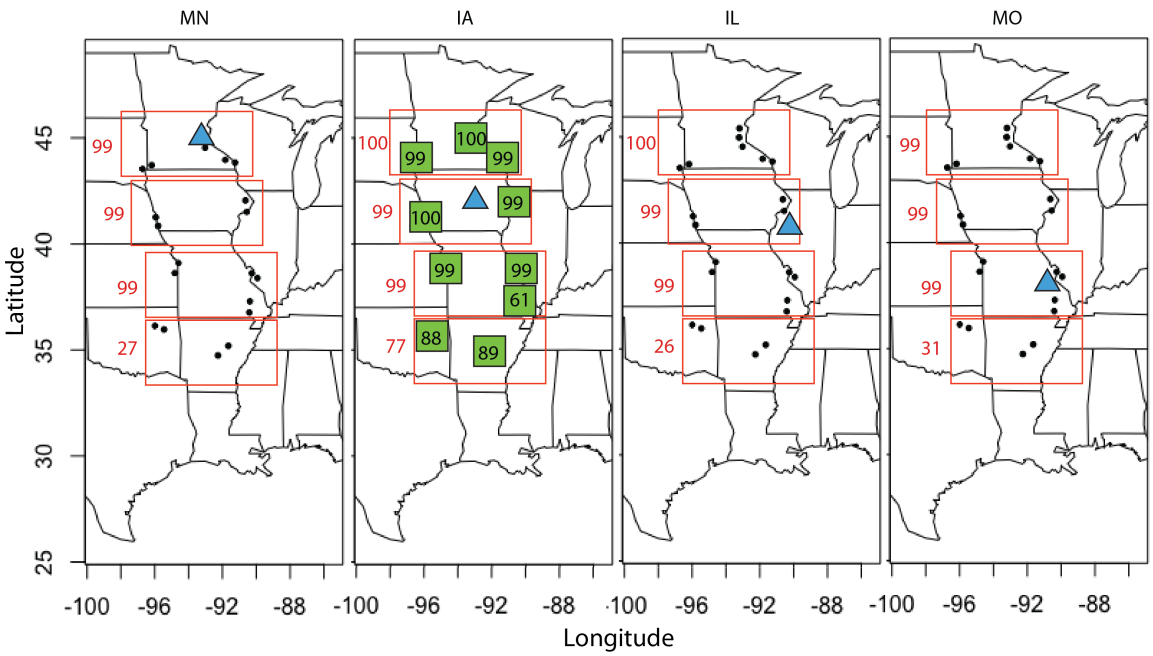

Figure 6. Mean probability of flowering across spatial scales, corresponding to the results in Table 6. Each red box and text delineate a separate hardiness zone and the associated mean probability of survival for individuals from that zone. Each green box is a regional grouping within a zone, and the number indicates the mean probability of survival for individuals from that region. Sites where 'region' did not have a significant effect on the probability of flowering do not have regional groups. Blue triangles indicate the locations of each site, and black points indicate the location of individual populations.

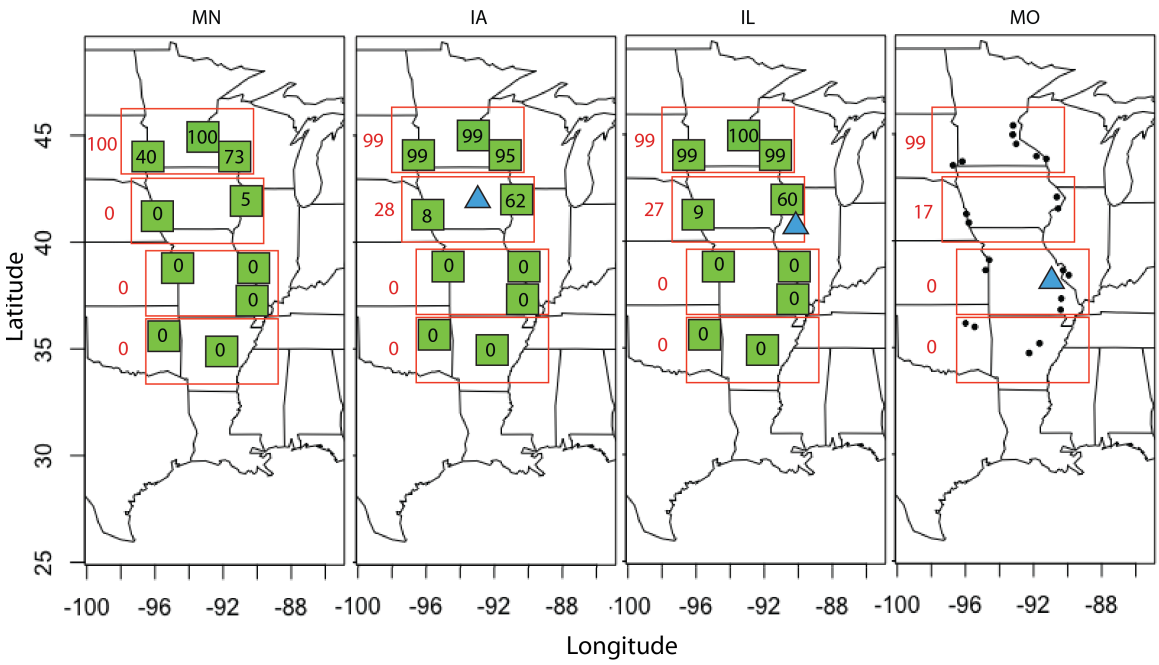

Figure 7. Mean probability of producing mature fruits by 16 weeks across spatial scales, corresponding to the results in Table 6. Each red box and text delineate a separate hardiness zone and the associated mean probability of survival for individuals from that zone. Each green box is a regional grouping within a zone, and the number indicates the mean probability of survival for individuals from that region. Sites where 'region' did not have a significant effect on the probability of flowering do not have regional groups. Blue triangles indicate the locations of each site, and black points indicate the location of individual populations. Note that for MN, the location of the site is masked by the regional groups.

Tables – Chapter 1

Table 1. Selection differentials from linear regression models.

		Transition to reproduction		First open male flower		Male flower to female flower		Height	
		Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural
Urban									
USC	S_{male}	-0.14	**-0.20**	**-0.13**[a]	**-0.32**[a]	**0.18**	**0.28**	**0.33**	**0.48**
	S_{female}	-0.09[a]	**-0.30**[a]	-0.05	**-0.23**	-0.05[a]	**0.18**[a]	**0.29**	**0.41**
USP	S_{male}	-0.15	0.04	-0.06	-0.18	0.17	0.02	**0.38**	0.19
	S_{female}	-0.14	0.05	-0.05	-0.20	0.10	-0.06	**0.31**	0.14
Rural									
RRO	S_{male}	-0.12	**-0.18**	**-0.23**	**-0.15**	**0.36**	**0.23**	**0.40**	**0.24**
	S_{female}	**-0.28**	-0.10	**-0.22**	-0.07	**0.25**	0.09	**0.30**	**0.15**
RRT	S_{male}	-0.12	-0.18	-0.22[a]	0.10[a]	0.13	-0.06	**0.40**	0.28
	S_{female}	-0.24	-0.27	**-0.38**[a]	-0.07[a]	-0.07	0.02	**0.28**	**0.34**

Note: Differentials were estimated separately urban and rural genotypes

S = linear regression coefficient

Selection differentials in bold are significant at $p < 0.05$.

[a] Selection differentials which are significantly different between urban and rural seed sources at $p<0.05$

Table 2. Selection gradients from multiple linear regression models.

		Transition to reproduction		First open male flower		Male flower to female flower	
		Urban	Rural	Urban	Rural	Urban	Rural
Urban							
USC	β_{male}	-0.06	0.07	-0.02	**-0.25**	0.13	0.14
	β_{female}	-0.09	-0.06	-0.08	-0.18	-0.14	0.01
USP	β_{male}	-0.31	0.14	0.34	-0.25	0.29[b]	-0.03[b]
	β_{female}	-0.28	0.11	0.27	-0.33	0.19	-0.17
Rural							
RRO	β_{male}	-0.06	-0.10	-0.01	0.11	**0.32**	**0.26**
	β_{female}	-0.12	-0.09	-0.06	0.05	0.16	0.09
RRT	β_{male}	-0.14[b]	**-0.42[b]**	-0.29[a]	0.21[a]	0.03	-0.09
	β_{female}	0.01	**-0.43**	**-0.54[a]**	0.03[a]	**-0.40**	-0.11

Note: Gradients were estimated separately urban and rural genotypes

β = linear multiple regression coefficient.

Selection gradients in bold are significant at $p < 0.05$.

[a] Selection gradients which are significantly different between urban and rural seed sources at $p<0.05$

[b] Selection gradients which are different between urban and rural seed sources at $p <0.10$

Tables – Chapter 3

Table 1. Description of populations that were sequenced. Populations are listed in ascending order by latitude. *n* indicates the number of maternal families that were sequenced from that population.

Population	Latitude	Longitude	*n*
LA4	29.9626	-90.1095	5
LA1	29.9815	-90.02892	5
BBN	30.36979	-91.10567	5
LDV	35.19796198	-91.64370698	5
WAG	35.96535395	-95.45782392	5
CLW	37.29567794	-90.36880106	5
SR9	38.38645	-89.92686	5
SU8	38.61736	-90.26415	5
PAO	38.62834563	-94.81757954	5
MPL	40.84571607	-95.80212162	5
PRC	42.05130841	-90.63232642	5
BMO	43.71449138	-96.18077593	5
LEW	43.96234593	-91.83442563	4
R10	44.53656	-93.02021	5
U25	44.96804	-93.22023	5

Table 2. Genetic diversity estimates for the study populations. Populations are listed in order of decreasing latitude.

	π	θ_w	Tajima's *D*
U25	2.06	2.25	-0.74
R10	2.02	2.18	-0.61
LEW	2.05	2.19	-0.77
BMO	2.08	2.27	-0.71
PRC	2.06	2.26	-0.74
MPL	2.06	2.25	-0.74
PAO	2.05	2.22	-0.66
SU8	2.07	2.26	-0.74
SR9	2.03	2.22	-0.71
CLW	2.07	2.28	-0.76
WAG	2.00	2.20	-0.73
LDV	2.02	2.18	-0.63
BBN	1.94	2.05	-0.47
LA1	1.99	2.13	-0.59
LA4	1.88	1.94	-0.33

π = nucleotide diversity ($\times 10^{-3}$). θ_w = Watterson estimator ($\times 10^{-2}$)

Table 3. Pairwise F_{ST} values among the sampled populations

	U25	R10	LEWI	BLMO	PRCR	MCPL	PAO	SU8	SR9	CLD	WAG	LDV	BBN	LA1
R10	0.07													
LEWI	0.07	0.07												
BLMO	0.07	0.07	0.07											
PRCR	0.06	0.07	0.06	0.07										
MCPL	0.07	0.07	0.08	0.07	0.07									
PAO	0.08	0.08	0.08	0.08	0.07	0.06								
SU8	0.06	0.07	0.07	0.07	0.06	0.07	0.08							
SR9	0.07	0.07	0.07	0.07	0.06	0.07	0.08	0.06						
CLD	0.06	0.07	0.06	0.07	0.06	0.07	0.07	0.06	0.06					
WAG	0.07	0.08	0.08	0.08	0.07	0.07	0.07	0.07	0.08	0.07				
LDV	0.09	0.08	0.08	0.08	0.07	0.08	0.09	0.07	0.07	0.06	0.08			
BBN	0.09	0.11	0.10	0.10	0.09	0.11	0.11	0.09	0.09	0.08	0.10	0.08		
LA1	0.09	0.10	0.09	0.09	0.09	0.10	0.11	0.08	0.09	0.08	0.10	0.08	0.08	
LA4	0.11	0.12	0.12	0.12	0.11	0.13	0.13	0.11	0.11	0.10	0.12	0.10	0.11	0.09

Tables – Chapter 4

Table 1. Description of populations included in the latitudinal reciprocal transplant experiment. Populations are listed in ascending order by latitude. *n* indicates the number of maternal plants that were included from that population. 'Climate zone' was determined by the USDA hardiness zones. 'Region' is nested within climate zone. Populations that are geographically nearby each other belong the same region.

Population	Latitude	Longitude	*n*	Climate zone	Region	Fruit counts
LA4	29.9626	-90.1095	24	9	9x	Y
LA6	29.97096	-90.01688	16	9	9x	
LA1	29.9815	-90.02892	18	9	9x	
BBN	30.36979	-91.10567	24	9	9x	
LTR	34.74136566	-92.25947634	23	7	7x	Y
LDV	35.19796198	-91.64370698	25	7	7x	
WAG	35.96535395	-95.45782392	25	7	7y	
ARK	36.13697554	-95.99441688	19	7	7y	
POP	36.76623098	-90.40547492	15	6	6y	
CLW	37.29567794	-90.36880106	25	6	6y	Y
SR9	38.38645	-89.92686	25	6	6x	
SU8	38.61736	-90.26415	25	6	6x	Y
PAO	38.62834563	-94.81757954	25	6	6z	
KAC	39.10268604	-94.60068525	25	6	6z	
MPL	40.84571607	-95.80212162	25	5	5y	
OMA	41.25688838	-95.94715796	12	5	5y	
ROI	41.51026297	-90.56311849	23	5	5x	Y
DMS	41.57788835	-93.62835764	25	5	5z	
PRC	42.05130841	-90.63232642	25	5	5x	Y
SFA	43.53882148	-96.72653811	24	4	4z	
BMO	43.71449138	-96.18077593	25	4	4z	
LAC	43.83766655	-91.2513499	25	4	4y	Y
LEW	43.96234593	-91.83442563	18	4	4y	
R10	44.53656	-93.02021	25	4	4x	Y
U25	44.96804	-93.22023	18	4	4x	Y
R16	45.40264	-93.19718	25	4	4x	Y

Table 2. Vegetative and reproductive scoring that was used on each individual plant at weeks 4,8,12 and 16. Note that an individual could receive an intermediate score between stages. The female flowering stage 4 was used as a cut-off for the presence of mature fruit at 16 weeks.

Stage	Overall plant	Male flowering	Female flowering
1	Vegetative; only leaves	Male inflorescences developing, only buds visible	Few stigmas visible
2	Initiation of reproduction	Male inflorescences visible, but flowers closed	Many stigmas visible, few immature fruit
3	Male or female flowers open	<50% flowers open per inflorescence	Many immature fruit visible, few stigmas
4	Fruit visible	>50% flowers open per inflorescence	Many immature fruit visible, few mature fruit
5	Senesced	All male flowers open or flowers senesced	Many mature fruit or senesced

Table 3. The effect of seed source latitude on whole plant reproductive stage at 12 weeks. Latitude (fixed) and block (random) were included as predictors. Fixed effects were evaluated with type II Wald chi-square test, while random effects were evaluated with LRT comparing models with and without block (LRT df = 1). Residual deviance and df were obtained from the model without block. The results of these models correspond to Figure 2.

Site	df	χ^2 or deviance	P-value
MN			
Latitude	1	1292	**<0.0001**
Block		296.4	<0.0001
Residual deviance	554	127	
IA			
Latitude	1	1246	**<0.0001**
Block		321.5	**<0.0001**
Residual deviance	525	137.1	
IL			
Latitude	1	840.3	**<0.0001**
Block		229.8	**<0.0001**
Residual deviance	367	100.8	
MO			
Latitude	1	1370.8	**<0.0001**
Block		333.7	<0.0001
Residual deviance	482	117.3	

Table 4. The effect of an individual's geographic distance from field site on conditional fitness components. Geographic distance (fixed) and block (random) were included as predictors. Fixed effects were evaluated with type II Wald chi-square test, while random effects were evaluated with LRT comparing models with and without block (LRT df = 1). Residual deviance and df were obtained from the model without block. The results of these models correspond to Figure 3.

Fitness components	Factor	df	χ^2 or deviance	P-value
MN				
P(Survival)	Geographic distance	1	6.5	**0.01**
	Block		0	1
	Residual deviance	647	497	
P(Flowering)	Geographic distance	1	82.3	**<0.0001**
	Block		0.09	0.76
	Residual deviance	562	264	
P(Mature fruit)	Geographic distance	1	85.4	**<0.0001**
	Block		0	1
	Residual deviance	408	156	
Number of fruits – 16 wks	Geographic distance	1	107.9	**<0.0001**
	Block		6756	**<0.0001**
	Residual deviance	54	9375	
IA				
P(Survival)	Geographic distance	1	0.6	0.42
	Block		0.8	0.35
	Residual deviance	647	605	
P(Flowering)	Geographic distance	1	129	**<0.0001**
	Block		0	1
	Residual deviance	532	302	
P(Mature fruit)	Geographic distance	1	42.4	**<0.0001**
	Block		0	1
	Residual deviance	412	501	
Number of fruits – 16 wks	Geographic distance	1	1331	**<0.0001**
	Block		13309	**<0.0001**
	Residual deviance	70	33045	
IL				
P(Survival)	Geographic distance	1	28.3	**<0.0001**
	Block		17.7	**<0.0001**
	Residual deviance	645	854	
P(Flowering)	Geographic distance	1	49.5	**<0.0001**
	Block		0	1
	Residual deviance	359	214	
P(Mature fruit)	Geographic distance	1	22	**<0.0001**
	Block		0	1
	Residual deviance	291	374	
Number of fruits – 16 wks	Geographic distance	1	5904	**<0.0001**
	Block		8311	**<0.0001**
	Residual deviance	64	23525	

MO				
P(Survival)	Geographic distance	1	0.20	0.65
	Block		0	1
	Residual deviance	634	660	
P(Flowering)	Geographic distance	1	75.3	**<0.0001**
	Block		0	1
	Residual deviance	498	542	
P(Mature fruit)	Geographic distance	1	86.4	**<0.0001**
	Block		0	1
	Residual deviance	331	243	
Number of fruits – 16 wks	Geographic distance	1	205	**<0.0001**
	Block		6421	**<0.0001**
	Residual deviance	48	9614	

Table 5. The effect of an individual's latitudinal distance from field site on conditional fitness components. Latitude distance (fixed) and block (random) were included as predictors. Fixed effects were evaluated with type II Wald chi-square test, while random effects were evaluated with LRT comparing models with and without block (LRT df = 1). Residual deviance and df were obtained from the model without block. The results of these models correspond to Figure 4.

Fitness components	Factor	df	χ^2 or deviance	P-value
MN				
P(Survival)	Latitude distance	1	3.82	**0.05**
	Block		0	1
	Residual deviance	647	500	
P(Flowering)	Latitude distance	1	84.2	**<0.0001**
	Block		0.09	0.76
	Residual deviance	562	266.9	
P(Mature fruit)	Latitude distance	1	50.4	**<0.0001**
	Block		0	1
	Residual deviance	408	136.6	
Number of fruits – 16 wks	Latitude distance	1	78.9	**<0.0001**
	Block		6885	**<0.0001**
	Residual deviance	54	9482	
IA				
P(Survival)	Latitude distance	1	1.3	0.26
	Block		0.8	0.35
	Residual deviance	647	605.1	
P(Flowering)	Latitude distance	1	123.5	**<0.0001**
	Block		0	1
	Residual deviance	532	326.7	
P(Mature fruit)	Latitude distance	1	111.3	**<0.0001**
	Block		1	0
	Residual deviance	412	250.8	
Number of fruits – 16 wks	Latitude distance	1	2000	**<0.0001**
	Block		13310	**<0.0001**
	Residual deviance	70	32409	
IL				
P(Survival)	Latitude distance	1	13.3	**0.0002**
	Block		16.1	**<0.0001**
	Residual deviance	645	870.1	
P(Flowering)	Latitude distance	1	50.8	**<0.0001**
	Block		0	1
	Residual deviance	369	174.5	
P(Mature fruit)	Latitude distance	1	78.2	**<0.0001**
	Block		0	1
	Residual deviance	291	183.1	
Number of fruits – 16 wks	Latitude distance	1	5939	**<0.0001**
	Block		8296	**<0.0001**
	Residual deviance	64	23480	

MO				
P(Survival)	Latitude distance	1	3.2	0.07
	Block		0.1	0.83
	Residual deviance	634	656.9	
P(Flowering)	Latitude distance	1	115.9	**<0.0001**
	Block		0.2	0.67
	Residual deviance	498	351.5	
P(Mature fruit)	Latitude distance	1	81.1	**<0.0001**
	Block		0	1
	Residual deviance	331	167.1	
Number of fruits – 16 wks	Latitude distance	1	185.3	**<0.0001**
	Block		6406	**<0.0001**
	Residual deviance	48	9619	

Table 6. The effect of spatial scale on conditional fitness components. 'Zone' refers to the USDA hardiness zone for the location of each population, 'region' refers to the spatial grouping within a given hardiness zone, 'population' refers to the individual populations found within a given regional grouping. All predictors were evaluated using a likelihood ratio test comparing nested models. The results of these models correspond to Figure 5,6,7.

Fitness components	Factor	df	Deviance	P-value
MN				
P(Survival)	Climate zone	3	11.47	**0.009**
	Region(Climate zone)	6	44.26	**<0.0001**
	Population(Region(Climate zone))	11	57.25	**<0.0001**
	Residual deviance	503	251.67	
P(Flowering)	Climate zone	3	203.36	**<0.0001**
	Region(Climate zone)	6	13.59	0.26
	Residual deviance	456	223.99	
P(Mature fruit)	Climate zone	3	301.69	**<0.0001**
	Region(Climate zone)	7	24.18	**0.001**
	Population(Region(Climate zone))	11	13.29	0.27
	Residual deviance	364	127.59	
IA				
P(Survival)	Climate zone	3	8.35	**0.04**
	Region(Climate zone)	6	28.88	**<0.0001**
	Population(Region(Climate zone))	11	33.48	**0.0004**
	Residual deviance	504	419.68	
P(Flowering)	Climate zone	3	29.24	**<0.0001**
	Region(Climate zone)	6	12.14	**0.06**
	Population(Region(Climate zone))	14	34.78	**0.001**
	Residual deviance	408	194.73	
P(Mature fruit)	Climate zone	3	355.41	**<0.0001**
	Region(Climate zone)	6	30.17	**<0.0001**
	Population(Region(Climate zone))	12	15.65	0.21
	Residual deviance	368	123.72	
IL				
P(Survival)	Climate zone	3	6,28	0.1
	Region(Climate zone)	6	35.05	**<0.0001**
	Population(Region(Climate zone))	11	22.37	**0.02**
	Residual deviance	503	633.9	
P(Flowering)	Climate zone	3	135.23	**<0.0001**
	Region(Climate zone)	6	4.16	0.65
	Population(Region(Climate zone))	11	14.72	0.19
	Residual deviance	303	114.18	
P(Mature fruit)	Climate zone	3	234.06	**<0.0001**
	Region(Climate zone)	6	25.74	**0.0002**
	Population(Region(Climate zone))	12	27.15	**0.007**
	Residual deviance	255	91.54	
MO				
P(Survival)	Climate zone	3	15.78	**0.001**
	Region(Climate zone)	6	30.35	**<0.0001**
	Population(Region(Climate zone))	11	39.28	**<0.0001**

	Residual deviance	491	455.06	
P(Flowering)	Climate zone	3	133.98	**<0.0001**
	Region(Climate zone)	6	6.17	0.40
	Population(Region(Climate zone))	11	22.42	**0.02**
	Residual deviance	378	245.45	
P(Mature fruit)	Climate zone	3	268.47	**<0.0001**
	Region(Climate zone)	6	1.81	0.94
	Population(Region(Climate zone))	12	3.89	0.98
	Residual deviance	292	139.46	

Bibliography

Agren, J., and D. W. Schemske. 2012. Reciprocal transplants demonstrate strong adaptive differentiation of the model organism *Arabidopsis thaliana* in its native range. New Phytol. 194:1112–1122.

Aitken, S. N., and M. C. Whitlock. 2012. Assisted gene flow to facilitate local adaptation to climate change. Annual Review of Ecology Evolution and Systematics 44:131010123440000.

Aitken, S. N., S. Yeaman, J. A. Holliday, T. Wang, and S. Curtis-McLane. 2008. Adaptation, migration or extirpation: climate change outcomes for tree populations. Evolutionary Applications 1:95–111.

Alberti, M. 2015. Eco-evolutionary dynamics in an urbanizing planet. Trends in Ecology and Evolution 30:114–126.

Andrews, K. R., J. M. Good, M. R. Miller, G. Luikart, and P. A. Hohenlohe. 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. Nature Reviews Genetics 17:81–92.

Antonovics, J., and A. D. Bradshaw. 1970. Evolution in closely adjacent plant populations. VIII. Clinal patterns at a mine boundary. Heredity 25: 349 - 362.

Arnold, B., R. B. Corbett-Detig, D. Hartl, and K. Bomblies. 2013. RADseq underestimates diversity and introduces genealogical biases due to nonrandom haplotype sampling. Molecular Ecology 22: 3179 - 2190.

Aronson, M. F. J., F. A. La Sorte, C. H. Nilon, M. Katti, M. A. Goddard, C. A. Lepczyk, P. S. Warren, N. S. G. Williams, S. Cilliers, B. Clarkson, C. Dobbs, R. Dolan, M. Hedblom, S. Klotz, J. L. Kooijmans, I. Macgregor-fors, M. Mcdonnell, U. Mörtberg, P. Pysek, S. Siebert, P. Werner, M. Winter, S. G. Williams, J. Sushinsky, and P. Pys. 2014. A global analysis of the impacts of urbanization on bird and plant diversity reveals key anthropogenic drivers. Proceedings of the Royal Society B 281: 20133330

Atwell, J. W., G. C. Cardoso, D. J. Whittaker, S. Campbell-Nelson, K. W. Robertson, and E. D. Ketterson. 2012. Boldness behavior and stress physiology in a novel environment suggest rapid correlated evolutionary adaptation. Behavioural Ecology 23:960–969.

Baskin, J. M., and C. C. Baskin. 1980. Ecophysiology of secondary dormancy in seeds of *Ambrosia artemisiifolia*. Ecology 61:475–480.

Baskin, J. M., and C. C. Baskin. 1977. Role of temperature in the germination ecology of three summer weeds. Oecologia 382:377–382.

Bass, D. J., V. Delpech, J. Beard, P. Bass, and R. S. Walls. 2000. Ragweed in Australia. Aerobiologia (Bologna). 16:107–111.

Bassett, I. J., and C. W. Crompton. 1975. The biology of Canadian weeds. II. *Ambrosia artemisiifolia* L. and *A. psilostachya*. Canadian Journal of Plant Science 55:463–476.

Bates, D., M. Machler, B. Bolker, and S. Walker. 2015. Fitting linear mixed-effects models using lme4. Journal of Statistical Software 67:1–48.

Baudis, M., R. H. Ellerbrock, K. Felsmann, A. Gessler, K. Gimbel, Z. Kayler, H.

Puhlmann, A. Ulrich, M. Weiler, E. Welk, and H. Bruelheide. 2014. Intraspecific differences in responses to rainshelter-induced drought and competition of *Fagus sylvatica* L. across Germany. Forest Ecology Management 330:283–293.
Becker, U., G. Colling, P. Dostal, A. Jakobsson, and D. Matthies. 2006. Local adaptation in the monocarpic perennial *Carlina vulgaris* at different spatial scales across Europe. Oecologia 150:506–18.
Beier, C., C. Beierkuhnlein, T. Wohlgemuth, J. Penuelas, B. Emmett, C. K??rner, H. de Boeck, J. H. Christensen, S. Leuzinger, I. A. Janssens, and K. Hansen. 2012. Precipitation manipulation experiments - challenges and recommendations for the future. Ecology Letters 15:899–911.
Beierkuhnlein, C., D. Thiel, A. Jentsch, E. Willner, and J. Kreyling. 2011. Ecotypes of European grass species respond differently to warming and extreme drought. Journal of Ecology 99:703–713.
Berardini, T. Z., L. Reiser, D. Li, Y. Mezheritsky, R. Muller, E. Strait, and E. Huala. 2015. The arabidopsis information resource: Making and mining the "gold standard" annotated reference plant genome. Genesis 53:474–485.
Bischoff, A., L. Crémieux, M. Smilauerova, C. S. Lawson, S. R. Mortimer, J. Dolezal, V. Lanta, A. R. Edwards, A. J. Brook, M. Macel, J. Leps, T. Steinger, and H. Müller-Schärer. 2006. Detecting local adaptation in widespread grassland species: the importance of scale and local plant community. Journal of Ecology 94:1130–1142.
Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120.
Brachi, B., R. Villoutreix, N. Faure, N. Hautekèete, Y. Piquot, M. Pauwels, D. Roby, J. Cuguen, J. Bergelson, and F. Roux. 2013. Investigation of the geographical scale of adaptive phenological variation and its underlying genetics in *Arabidopsis thaliana*. Molecular Ecology 22:4222–4240.
Bradburd, G. S., P. L. Ralph, and G. M. Coop. 2013. Disentangling the effects of geographic and ecological isolation on genetic differentiation. Evolution 67:3258–3273.
Burghardt, L. T., C. J. E. Metcalf, A. M. Wilczek, J. Schmitt, and K. Donohue. 2014. Modeling the influence of genetic and environmental variation on the expression of plant life cycles across landscapes. American Naturalist 185:212–227.
Cadenasso, M. L., S. T. A. Pickett, and K. Schwarz. 2007. Spatial heterogeneity in urban ecosystems: Reconceptualizing land cover and a framework for classification. Frontiers in Ecology and the Environment 5:80–88.
Campbell, D. R., and C. Wendlandt. 2013. Altered precipitation affects plant hybrids differently than their parental species. American Journal of Botany 100:1322–1331.
Catchen, J., S. Bassham, T. Wilson, M. Currey, C. O'Brien, Q. Yeates, and W. A. Cresko. 2013. The population structure and recent colonization history of Oregon threespine stickleback determined using restriction-site associated DNA-sequencing. Molecular Ecology 22:2864–2883.
Catchen, J. M., P. A. Hohenlohe, L. Bernatchez, W. C. Funk, K. R. Andrews, and F. W. Allendorf. 2017. Unbroken: RADseq remains a powerful tool for understanding the genetics of adaptation in natural populations. Molecular Ecology Resources 17:362–365.
Cavender-Bares, J., A. González-Rodríguez, D. A. R. Eaton, A. A. L. Hipp, A. Beulke,

and P. S. Manos. 2015. Phylogeny and biogeography of the american live oaks (*Quercus* subsection *Virentes*): A genomic and population genetics approach. Molecular Ecology 24:3668–3687.
Cavender-Bares, J., and J. A. Ramírez-Valiente. 2017. Physiological evidence from common garden experiments for local adaptation and adaptive plasticity to climate in American live oaks (*Quercus* section V*irentes*): Implications for Conservation Under Global Change. Oaks Physiological Ecology. Exploring the Functional Diversity of Gene *Quercus*. Tree Physiology 7: 107–135.
Caye, K., B. Jumentier, J. Lepeule, and O. Franc. 2019. LFMM 2 : Fast and accurate inference of gene-environment associations in genome-wide studies. Molecular Biology and Evolution 36: 852-860.
Chambers, S. M., and N. C. Emery. 2016. Population differentiation and countergradient variation throughout the geographic range in the fern gametophyte *Vittaria appalachiana*. American Journal of Botany 103:86–98.
Chapman, D. S., T. Haynes, S. Beal, F. Essl, and J. M. Bullock. 2014. Phenology predicts the native and invasive range limits of common ragweed. Global Change Biology 20:192–202.
Chauvel, B., F. Dessaint, C. Cardinal-Legrand, and F. Bretagnolle. 2006. The historical spread of *Ambrosia artemisiifolia* L. in France from herbarium records. Journal of Biogeography 33:665–673.
Cheng, C., B. J. White, C. Kamdem, K. Mockaitis, C. Costantini, M. W. Hahn, and N. J. Besansky. 2012. Ecological genomics of *Anopheles gambiae* along a latitudinal cline: A population-resequencing approach. Genetics 190:1417–1432.
Cheptou, P.-O., O. Carrue, S. Rouifed, and A. Cantarel. 2008. Rapid evolution of seed dispersal in an urban environment in the weed *Crepis sancta*. Procedings of the National Academy of Science USA105:3796–9.
Clausen, J., D. D. Keck, and W. M. Hiesey. 1940. Experimental studies on the nature of species I. The effects of varied environments on western American plants. Carnegie Institute, Washington DC.
Colautti, R. I., and S. C. H. Barrett. 2013. Rapid adaptation to climate facilitates range expansion of an invasive plant. Science 342:364–366.
Combosch, D. J., and S. V. Vollmer. 2015. Trans-Pacific RAD-Seq population genomics confirms introgressive hybridization in Eastern Pacific Pocillopora corals. Molecular Phylogenetics and Evolution 88:154–162.
Conover, D. O., and T. M. C. Present. 1990. Countergradient variation in growth rate: compensation for length of the growing season among Atlantic silversides from different latitudes. Oecologia 83:316–324.
Conover, D., and E. T. Schultz. 1995. Phenotypic similarity and the evolutionary significance of countergradient variation. Trends in Ecology and Evolution 10: 248 - 252.
Cunningham, M. A., E. Snyder, D. Yonkin, M. Ross, and T. Elsen. 2007. Accumulation of deicing salts in soils in an urban environment. Urban Ecosystems 11:17–31.
Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, and R. Durbin. 2011. The variant call format and VCFtools. Bioinformatics 27:2156–2158.
Davis, M. B., and R. G. Shaw. 2001. Range shifts and adaptive responses to Quaternary

climate change. Science 292:673–679.
Deen, W., T. Hunt, and J. Swanton. 1998. Influence of temperature , photoperiod , and irradiance on the phenological development of common ragweed (*Ambrosia artemisiifolia*). Weed Science 46:555–560.
Didiano, T. J., M. T. J. Johnson, and T. P. Duval. 2016. Disentangling the effects of precipitation amount and frequency on the performance of 14 grassland species. PLoS One 11:1–19.
Donohue, K., R. Rubio de Casas, L. Burghardt, K. Kovach, and C. G. Willis. 2010. Germination, Postgermination Adaptation, and Species Ecological Ranges. Annual Review of Ecology Evolution and Systematics 41:293–319.
Dorn, L. A., and T. Mitchell-Olds. 2006. Genetics of *Brassica campestris*. 1. Genetic constraints on evolution of life-history characters. Evolution 45:371.
Dudley, S. A. 1996. Differing selection on plant physiological traits in response to environmental water availability: a test of adaptive hypotheses. Evolution 50:92–102.
Dunnett, N. P., A. J. Willis, R. Hunt, and J. P. Grime. 1998. A 38-year study of relations between weather and vegetation dynamics in road verges near Bibury, Gloucestershire. Journal of Ecology 86:610–623.
Eckhart, V. M., M. A. Geber, C. M. McGuire, and C. F. Williams. 2004. Experimental studies of adataption in *Clarkia xantiana*. I. Sources of trait variation across a subspecies border. Evolution 58:59–70.
Eisenhauer, N., S. Cesarz, R. Koller, K. Worm, and P. B. Reich. 2012. Global change belowground: Impacts of elevated CO 2, nitrogen, and summer drought on soil food webs and biodiversity. Global Change Biology 18:435–447.
Engelbrecht, B. M. J., L. S. Comita, R. Condit, T. A. Kursar, M. T. Tyree, B. L. Turner, and S. P. Hubbell. 2007. Drought sensitivity shapes species distribution patterns in tropical forests. Nature 447:80–82.
Epstein, B., M. Jones, R. Hamede, S. Hendricks, H. McCallum, E. P. Murchison, B. Schönfeld, C. Wiench, P. Hohenlohe, and A. Storfer. 2016. Rapid evolutionary response to a transmissible cancer in Tasmanian devils. Nature Communications 7.
Essl, F., K. Biró, D. Brandes, O. Broennimann, J. M. Bullock, D. S. Chapman, B. Chauvel, S. Dullinger, B. Fumanal, A. Guisan, G. Karrer, G. Kazinczi, C. Kueffer, B. Laitung, C. Lavoie, M. Leitner, T. Mang, D. Moser, H. Müller-Schärer, B. Petitpierre, R. Richter, U. Schaffner, M. Smith, U. Starfinger, R. Vautard, G. Vogl, M. von der Lippe, and S. Follak. 2015. Biological Flora of the British Isles: *Ambrosia artemisiifolia*. Journal of Ecology 103:1069–1098.
Fenger, J. 1999. Urban air quality. Atmospheric Environment 33:4877–4900.
Fenster, C. B., and L. F. Galloway. 2000. Population differentiation in an annual legume: genetic architecture. Evolution 54:1157–1172.
Fick, S. E., and R. J. Hijmans. 2017. Worldclim 2: new 1-km spatial resolution climate surfaces for global land areas. International Journal of Climatology 37:4302–4315.
Fischer, G., F. Nachtargaele, S. Prieler, H. T. van Velthuizen, L. Verelst, and D. Wiberg. 2008. Global agro-ecological zons assessment for agriculture. Lanxenburg, Austria and Rome, Italy.
Forester, B. R., J. R. Lasky, H. H. Wagner, and D. L. Urban. 2018. Comparing methods for detecting multilocus adaptation with multivariate genotype–environment

associations. Molecular Ecology. 27:2215–2233.
Fournier-Level, A., A. Korte, M. D. Cooper, M. Nordborg, J. Schmitt, and A. M. Wilczek. 2011. A map of local adaptation in *Arabidopsis thaliana*. Science 334:86–9.
Fox, J., and S. Weisberg. 2011. An {R} companion to applied regression. Thousand Oaks, CA.
Franks, S. J., S. Sim, and A. E. Weis. 2007. Rapid evolution of flowering time by an annual plant in response to a climate fluctuation. Proceedings of the National Academy of Sciences USA 104:1278–82.
Frenz, D. A. 2001. Interpreting atmospheric pollen counts for use in clinical allergy: allergic symptomology. Annals of Allergy, Asthma and Immunology 86:150–158.
Friedman, J., and S. C. H. Barrett. 2008. High outcrossing in the annual colonizing species *Ambrosia artemisiifolia* (Asteraceae). Annals of Botany 101:1303–9.
Fu, L., B. Niu, Z. Zhu, S. Wu, and W. Li. 2012. CD-HIT: Accelerated for clustering the next-generation sequencing data. Bioinformatics 28:3150–3152.
Galloway, L. F., and C. B. Fenster. 2000. Population differentiation in an annual legume: local adaptation. Evolution 54:1173–1181.
Gard, B., F. Bretagnolle, F. Dessaint, and B. Laitung. 2013. Invasive and native populations of common ragweed exhibit strong tolerance to foliar damage. Basic Applied Ecology 14:28–35.
Garrison, E., and M. Marth. 2016. Haplotype-based variant detection from short-read sequencing arXiv Prepr. arXiv:1207.3907.
Gaudet, P., and C. Dessimoz. 2017. Gene ontology: pitfalls, biases and remedies. P. *in* C. Dessimoz and N. Skunca, eds. The Gene Ontology Handbook. Methods in Molecular Biology. Humana Press, New York, NY.
Geber, M. A. 1990. The cost of meristem limitation in *Polygonum arenastrum*: negative genetic correlations between fecundity and growth. Evolution 44:799.
Geber, M. a, and V. M. Eckhart. 2005. Experimental studies of adaptation in *Clarkia xantiana*. II. Fitness variation across a subspecies border. Evolution 59:521–531.
Genton, B. J., P. M. Kotanen, P.-O. Cheptou, C. Adolphe, and J. A. Shykoff. 2005a. Enemy release but no evolutionary loss of defence in a plant invasion: an inter-continental reciprocal transplant experiment. Oecologia 146:404–414.
Genton, B. J., J. A. Shykoff, and T. Giraud. 2005b. High genetic diversity in French invasive populations of common ragweed, *Ambrosia artemisiifolia,* as a result of multiple sources of introduction. Molecular Ecology 14:4275–85.
Geyer, C. J., S. Wagenius, and R. G. Shaw. 2007. Aster models for life history analysis. Biometrika 94:415–426.
Gorton, A. J., D. A. Moeller, and P. Tiffin. 2018. Little plant, big city: a test of adaptation to urban environments in common ragweed (Ambrosia artemisiifolia). Proceedings of the Royal Society of London B 285: 20180968.
Gorton, A. J., P. Tiffin, and D. A. Moeller. 2019. Does adaptation to historical climate shape plant responses to future rainfall patterns? A rainfall manipulation experiment with common ragweed. Oecologia 190: 941
Gregg, J. W., C. G. Jones, and T. E. Dawson. 2003. Urbanization effects on tree growth in the vicinity of New York City. Nature 424:183–187.
Grime, J. P. 1977. Evidence for the existence of three primary strategies in plants and its

relevance to ecological and evolutionary theory. American Naturalist 111:1169–1194.
Grimm, E. C. 2001. Trends and paleoecological problems in the vegetation and climate history of the Northern Great Plains. Biology and Envionment: Proceedings of the Royal Irish Academy 101:47–64.
Gugger, P. F., S. J. Cokus, and V. L. Sork. 2016. Association of transcriptome-wide sequence variation with climate gradients in valley oak (*Quercus lobata*). Tree Genetics and Genomes 12: 15.
Haas, B. J., A. Papanicolaou, M. Yassour, M. Grabherr, P. D. Blood, J. Bowden, M. B. Couger, D. Eccles, B. Li, M. Lieber, M. D. Macmanes, M. Ott, J. Orvis, N. Pochet, F. Strozzi, N. Weeks, R. Westerman, T. William, C. N. Dewey, R. Henschel, R. D. Leduc, N. Friedman, and A. Regev. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nature Protocols 8:1494–1512.
Hall, M. C., and J. H. Willis. 2006. Divergent selection on flowering time contributes to local adaptation in *Mimulus guttatus* populations. Evolution 60:2466–2477.
Hancock, A. M., B. Brachi, N. Faure, M. W. Horton, L. B. Jarymowycz, F. G. Sperone, C. Toomajian, F. Roux, and J. Bergelson. 2011. Adaptation to climate across the *Arabidopsis thaliana* genome. Science 334:83–6.
Hereford, J. 2009. A quantitative survey of local adaptation and fitness trade-offs. American Naturalist 173:579–88.
Hoban, S., J. L. Kelley, K. E. Lotterhos, M. F. Antolin, G. Bradburd, D. B. Lowry, M. L. Poss, L. K. Reed, A. Storfer, and M. C. Whitlock. 2016. Finding the genomic basis of local adaptation: pitfalls, practical solutions, and future directions. American Naturalist 188: 379-397.
Hodgins, K. A., and L. Rieseberg. 2011. Genetic differentiation in life-history traits of introduced and native common ragweed (*Ambrosia artemisiifolia*) populations. Journal of Evolutionary Biology 24:2731–49.
Hoffmann, A. A., and C. M. Sgrò. 2011. Climate change and evolutionary adaptation. Nature 470:479–85.
Homer, C. G., J. A. Dewitz, L. Yang, S. Jin, P. Danielson, G. Zian, J. Coulston, N. D. Herold, J. D. Wickham, and K. Megown. 2015. Completion of the 2011 National Land Cover Database for the conterminous United States - Representing a decade of land cover change information. Photogrammetric Enginerring and Remote Sensing 81:345–354.
Hothorn, T., F. Bretz, and P. Westfall. 2008. Simultaneous inference in general parametric models. Biometrical Journal 50:346–363.
Huff, F. A., and S. A. Changnon. 1973. Precipitation Modification By Major Urban Areas. Bulletin of the American Meterological Society 54: 1220 - 1232.
IPCC. 2013. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge.
Johnson, M. T. J., and J. Munshi-South. 2017. Evolution of life in urban environments. Science 358: 8327.
Jonas, C. S., and M. A. Geber. 1999. Variation among populations of *Clarkia unguiculata* (Onagraceae) along altitudinal and latitudinal gradients. American

Journal of Botany 86:333–343.
Jones, K. L. 1936. Studies on *Ambrosia* I. The inheritance of floral types in the ragweed, *Ambrosia elatior* L. American Midland. Naturalist 17:673–699.
Joshi, J., B. Schmid, M. C. Caldeira, P. G. Dimitrakopoulos, J. Good, R. Harris, A. Hector, K. Huss-Danell, A. Jumpponen, S. Minns, C. P. H. Mulder, J. S. Pereira, S. Prinz, M. Scherer-Lorenzen, A. S. D. Siamantziouras, A. C. Terry, A. Y. Troumbis, and J. H. Lawton. 2001. Local adaptation enhances performance of common plant species. Ecology Letters. 4:536–544.
Jump, A. S., and J. Penuelas. 2005. Running to stand still: adaptation and the response of plants to rapid climate change. Ecology Letters. 8:1010–1020.
Kartesz, J. T. 2013. The Biota of North America Program (BONAP).
Kawecki, T. J., and D. Ebert. 2004. Conceptual issues in local adaptation. Ecology Letters 7:1225–1241.
Keller, S. R., R. Y. Soolanayakanahally, R. D. Guy, S. N. Silim, M. S. Olson, and P. Tiffin. 2011. Climate-driven local adaptation of ecophysiology and phenology in balsam poplar, *Populus balsamifera* L. (Salicaceae). American Journal of Botany 98:99–108.
Kimball, S., A. L. Angert, T. E. Huxman, and D. L. Venable. 2010. Contemporary climate change in the Sonoran Desert favors cold-adapted species. Global Change Biology 16:1555–1565.
Knapp, S., L. Dinsmore, C. Fissore, S. E. Hobbie, I. Jakobsdottir, J. Kattge, J. Y. King, S. Klotz, J. P. McFadden, and J. Cavender-Bares. 2012. Phylogenetic and functional characteristics of household yard floras and their changes along an urbanization gradient. Ecology Letters 93:83–98.
Koornneef, M., C. Alonso-Blanco, A. J. M. Peeters, and W. Soppe. 1998. Genetic Control of Flowering Time in *Arabidopsis*. Annual Review of Plant Physiology and Plant Molecular Biology 49:345–370.
Korneliussen, T., A. Albrechtsen, and R. Nielsen. 2014. ANGSD: Analysis of next generation sequencing data. BMC Bioinformatics 15:356.
Kotchoni, S. O., C. Kuhns, A. Ditzer, H. H. Kirch, and D. Bartels. 2006. Over-expression of different aldehyde dehydrogenase genes in *Arabidopsis thaliana* confers tolerance to abiotic stress and protects plants against lipid peroxidation and oxidative stress. Plant, Cell and Environment 29:1033–1048.
Kozak, G. M., R. S. Brennan, E. L. Berdan, R. C. Fuller, and A. Whitehead. 2014. Functional and population genomic divergence within and between two species of killifish adapted to different osmotic niches. Evolution 68:63–80.
Kozlowski, J. 1992. Optimal allocation of resources to growth and reproduction: Implications for age and size at maturity. Trends in Ecology Evolution 7:15–19.
Kramer, P. J. 1951. Causes of injury to plants resulting from flooding of the soil. Plant Physiology 26:722–36.
Lacey, E. P. 1988. Latitudinal variation in reproductive timing of a short-lived monocarp, *Daucus carota* (Apiaceae). Ecology 69:220–232.
Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters. Evolution 37:1210–1226.
Lefebvre, V., H. North, A. Frey, B. Sotta, M. Seo, M. Okamoto, E. Nambara, and A. Marion-Poll. 2006. Functional analysis of *Arabidopsis* NCED6 and NCED9 genes

indicates that ABA synthesized in the endosperm is involved in the induction of seed dormancy. Plant Journal 45:309–319.
Lehikoinen, A., P. Byholm, E. Ranta, P. Saurola, J. Valkama, E. Korpimäki, H. Pietiäinen, and H. Henttonen. 2009. Reproduction of the common buzzard at its northern range margin under climatic change. Oikos 118:829–836.
Leiblein-Wild, M. C., and O. Tackenberg. 2014. Phenotypic variation of 38 European *Ambrosia artemisiifolia* populations measured in a common garden experiment. Biological Invasions 16:2003–2015.
Leimu, R., and M. Fischer. 2008. A meta-analysis of local adaptation in plants. PLoS One 3:e4010.
Lenormand, T. 2002. Gene flow and the limits to natural selection. Trends in Ecology and Evolution 17:183–189.
Lenth, R. V. 2016. Least-square means: the R package lsmeans. J. Stat. Softw. 69:1–33.
Levins, R. 1968. Evolution in changing environments: some theoretical explorations. Princeton University Press.
Levins, R. 1969. Thermal acclimation and heat resistance in Drosophila species. American Naturalist 103:483–499.
Lewis, W. H., P. Vinay, and W. E. Zenger. 1983. Airborne and allergnic pollen of North America. John Hopkins University Press, Baltimore, MD.
Li, C., O. Junttila, A. Ernstsen, P. Heino, and E. T. Palva. 2003. Photoperiodic control of growth, cold acclimation and dormancy development in silver birch (*Betula pendula*) ecotypes. Physiologia Plantarum 117:206–212.
Li, H., and R. Durbin. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26:589–595.
Li, X.-M., D.-Y. She, D.-Y. Zhang, and W.-J. Liao. 2014. Life history trait differentiation and local adaptation in invasive populations of *Ambrosia artemisiifolia* in China. Oecologia 177: 669-677.
Liancourt, P., L. A. Spence, D. S. Song, A. Lkhagva, A. Sharkhuu, B. Boldgiv, B. R. Helliker, P. S. Petraitis, and B. B. Casper. 2013. Plant response to climate change varies with topography, interactions with neighbors, and ecotype. Ecology 94:444–453.
Linhart, Y. B., and M. C. Grant. 1996. Evolutionary significance of local genetic differentiation in plants. Annual Review of Ecoogy and Systematics 27:237–277.
Lorenzo, C., M. Marco, D. M. Paola, C. Alfonso, O. Marzia, and O. Simone. 2006. Long distance transport of ragweed pollen as a potential cause of allergy in central Italy. Annals of Allergy, Asthma and Immunology 96:86–91.
Lotterhos, K. E., and M. C. Whitlock. 2015. The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. Molecular Ecology 24:1031–1046.
Lourenço, A., D. Álvarez, I. J. Wang, and G. Velo-Antón. 2017. Trapped within the city: integrating demography, time since isolation and population-specific traits to assess the genetic effects of urbanization. Molecular Ecology 26:1498–1514.
Lowry, D. B., S. Hoban, J. L. Kelley, K. E. Lotterhos, L. K. Reed, M. F. Antolin, and A. Storfer. 2017. Breaking RAD: an evaluation of the utility of restriction site-associated DNA sequencing for genome scans of adaptation. Molecular Ecology Resources 17:142–152.

Lowry, D. B., J. Schmutz, K. W. Barry, T. L. Logan, E. Meyer, J. A. Chapman, S. H. Taylor, T. E. Juenger, K. Hernandez, and D. S. Rokhsar. 2014. The genetics of divergence and reproductive isolation between ecotypes of *Panicum hallii* . New Phytol. 205:402–414.

Luquez, V., D. Hall, B. R. Albrectsen, J. Karlsson, P. Ingvarsson, and S. Jansson. 2008. Natural phenological variation in aspen (*Populus tremula*): The SwAsp collection. Tree Genetics Genomes 4:279–292.

MacDonald, A. A. M., and P. M. Kotanen. 2010a. Leaf damage has weak effects on growth and fecundity of common ragweed (*Ambrosia artemisiifolia*). Botany 88:158–164.

MacDonald, A. A. M., and P. M. Kotanen. 2010b. The effects of disturbance and enemy exclusion on performance of an invasive species, common ragweed, in its native range. Oecologia 162:977–86.

Marks, P. L. 1983. On the origin of the field plants of the northeastern United States. American Naturalist 122:210–228.

Martin, M. D., M. T. Olsen, J. A. Samaniego, E. A. Zimmer, and M. T. P. Gilbert. 2016. The population genomic basis of geographic differentiation in North American common ragweed (*Ambrosia artemisiifolia* L.). Ecology and Evolution 1–12.

Martin, M. D., E. a Zimmer, M. T. Olsen, A. D. Foote, M. T. P. Gilbert, and G. S. Brush. 2014. Herbarium specimens reveal a historical shift in phylogeographic structure of common ragweed during native range disturbance. Molecular Ecology 23:1701–16.

McKinney, G. J., W. A. Larson, L. W. Seeb, and J. E. Seeb. 2017. RADseq provides unprecedented insights into molecular ecology and evolutionary genetics: comment on Breaking RAD by Lowry et al. (2016). Molecular Ecology Resources 17:356–361.

McKinney, M. L. 2008. Effects of urbanization on species richness: A review of plants and animals. Urban Ecosystems 11:161–176.

Meisner, J., and A. Albrechtsen. 2018. Inferring population structure and admixture proportions in low-depth NGS data. Genetics 210:719–731.

Méndez-Vigo, B., F. X. Picó, M. Ramiro, J. M. Martínez-Zapater, and C. Alonso-Blanco. 2011. Altitudinal and climatic adaptation is mediated by flowering traits and FRI , FLC , and PHYC Genes in *Arabidopsis*. Plant Physiology 157:1942–1955.

Michaels, S. D., and R. M. Amasino. 2007. FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. Plant Cell 11:949.

Michaels, S. D., I. C. Bezerra, and R. M. Amasino. 2004. FRIGIDA-related genes are required for the winter-annual habit in *Arabidopsis.* Proceedings of the National Academy of Sciences 101:3281–5.

Milakovic, I., K. Fiedler, and G. Karrer. 2014. Management of roadside populations of invasive *Ambrosia artemisiifolia* by mowing. Weed Research 54:256–264.

Montalvo, A. M., and N. C. Ellstrand. 2000. Transplantation of the subshrub Lotus scoparious: testing the home-site advantage hypothesis. Conservation Biology 14:1034–1045.

Munshi-South, J., and K. Kharchenko. 2010. Rapid, pervasive genetic differentiation of urban white-footed mouse (*Peromyscus leucopus*) populations in New York City. Molecular Ecology 19:4242–4254.

Mustilli, A., S. Merlot, A. Vavasseur, F. Fenzi, and J. Giraudat. 2002. *Arabidopsis* OST1

protein kinase mediates the regulation of stomatal aperture by abscisic acid and acts upstream of reaction oxygen species production. The Plant Cell 14: 3089–3099.
Nacci, D. E., D. Champlin, and S. Jayaraman. 2010. Adaptation of the estuarine fish *Fundulus heteroclitus* (Atlantic Killifish) to polychlorinated biphenyls (PCBs). Estuaries and Coasts 33:853–864.
Neil, K., and J. Wu. 2006. Effects of urbanization on plant flowering phenology: A review. Urban Ecosystems 9:243–257.
NOAA. 2014. National Climate Assessment.
Novembre, J., and A. Di Rienzo. 2009. Spatial patterns of variation due to natural selection in humans. Nature Reviews Genetics 10:745–755.
Noy-Meir, I. 1973. Desert ecosystems: Environment and producers. Annual Review in Ecology and Systematics 4:23–51.
Oke, T. R. 1973. City size and the urban heat island. Atmos. Environ. Pergamon Pres 7:769–779.
Olsson, K., and J. Ågren. 2002. Latitudinal population differentiation in phenology, life history and flower morphology in the perennial herb *Lythrum salicaria*. Journal of Evolutionary Biology 15:983–996.
Patracchini, C., F. Vidotto, and A. Ferrero. 2011. Common ragweed (*Ambrosia artemisiifolia)* growth as affected by plant density and clipping. Weed Technology 25:268–276.
Pickett, S. T. A., M. L. Cadenasso, D. L. Childers, M. J. McDonnell, and W. Zhou. 2016. Evolution and future of urban ecological science: ecology *in* , *of* , and *for* the city. Ecosystem Health and Sustainability 2:e01229.
Pickett, S. T. A, M. L. Cadenasso, J. M. Grove, C. G. Boone, P. M. Groffman, E. Irwin, S. S. Kaushal, V. Marshall, B. P. McGrath, C. H. Nilon, R. V Pouyat, K. Szlavecz, A. Troy, and P. Warren. 2011. Urban ecological systems: scientific foundations and a decade of progress. Journal of Environmental Management 92:331–62.
Poorter, L., and F. Bongers. 2006. Leaf traits are good predictors of plant performance across 53 rain forest species. Ecology 87: 1733 - 1743
Qi, Z. C., Y. Yu, X. Liu, A. Pais, T. Ranney, R. Whetten, and Q. Y. Xiang. 2015. Phylogenomics of polyploid *Fothergilla* (Hamamelidaceae) by RAD-tag based GBS - Insights into species origin and effects of software pipelines. Journal of Systematics and Evolution 53:432–447.
Raabová, J., Z. Münzbergová, and M. Fischer. 2011. The role of spatial scale and soil for local adaptation in *Inula hirta*. Basic and Applied Ecology 12:152–160.
Ramírez-Valiente, J. A., and J. Cavender-Bares. 2017. Evolutionary trade-offs between drought resistance mechanisms across a precipitation gradient in a seasonally dry tropical oak (*Quercus oleoides*). Tree Physiology 37:889–901.
Ramírez-Valiente, J. A., N. J. Deacon, J. Etterson, A. Center, J. P. Sparks, K. L. Sparks, T. Longwell, G. Pilz, and J. Cavender-Bares. 2018. Natural selection and neutral evolutionary processes contribute to genetic divergence in leaf traits across a precipitation gradient in the tropical oak *Quercus oleoides*. Molecular Ecology 27:2176–2192. J
Rausher, M. D. 1992. The measurement of selection on quantitative traits: biases due to environmental covariances between traits and fitness. Evolution 46:616–626.
Raynor, G. S., E. C. Ogden, and J. V Hayes. 1970. Dispersion and deposition of ragweed

pollen from experimental sources. Journal of Meteorology 9.
Reekie, E. B., and B. F. A. 1987. Reproductive effort in plants. 3. Effect of reproduction on vegetative activity. American Naturalist 129:907919.
Reich, P. B., S. E. Hobbie, and T. D. Lee. 2014. Plant growth enhancement by elevated CO2 eliminated by joint water and nitrogen limitation. Nature Geosciences7:920–924.
Reich, P. B., C. Uhl, M. Walters, and D. Ellsworth. 1991. Leaf life-span as a determinant of leaf structure and function among 23 Amazonian tree species. Oecologia 86:16–24.
Reid, N. M., D. A. Proestou, B. W. Clark, W. C. Warren, J. K. Colbourne, J. R. Shaw, S. I. Karchner, M. E. Hahn, D. Nacci, M. F. Oleksiak, D. L. Crawford, and A. Whitehead. 2016. The genomic landscape of rapid repeated evolutionary adaptation to toxic pollution in wild fish. Science 354:1305 LP – 1308.
Rellstab, C., F. Gugerli, A. J. Eckert, A. M. Hancock, and R. Holderegger. 2015. A practical guide to environmental association analysis in landscape genomics. Molecular Ecology 24: 4348-4370.
Rockman, M. V. 2012. The QTN program and the alleles that matter for evolution: All that's gold does not glitter. Evolution 66:1–17.
Rubio de Casas, R., C. G. Willis, W. Pearse, C. Baskin, J. Baskin, and J. Cavendar-Barres. 2017. Global biogeography of seed dormancy is determined by seasonality and seed size in the legume. New Phytologist 214: 1527-1536.
Sakuma, Y., K. Maruyama, F. Qin, Y. Osakabe, K. Shinozaki, and K. Yamaguchi-Shinozaki. 2006. Dual function of an *Arabidopsis* transcription factor DREB2A in water-stress-responsive and heat-stress-responsive gene expression. Proceedings of the National Academy of Sciences USA 103:18822–18827.
Salojärvi, J., O. P. Smolander, K. Nieminen, S. Rajaraman, O. Safronov, P. Safdari, A. Lamminmäki, J. Immanen, T. Lan, J. Tanskanen, P. Rastas, A. Amiryousefi, B. Jayaprakash, J. I. Kammonen, R. Hagqvist, G. Eswaran, V. H. Ahonen, J. A. Serra, F. O. Asiegbu, J. De Dios Barajas-Lopez, D. Blande, O. Blokhina, T. Blomster, S. Broholm, M. Brosché, F. Cui, C. Dardick, S. E. Ehonen, P. Elomaa, S. Escamez, K. V. Fagerstedt, H. Fujii, A. Gauthier, P. J. Gollan, P. Halimaa, P. I. Heino, K. Himanen, C. Hollender, S. Kangasjärvi, L. Kauppinen, C. T. Kelleher, S. Kontunen-Soppela, J. P. Koskinen, A. Kovalchuk, S. O. Kärenlampi, A. K. Kärkönen, K. J. Lim, J. Leppälä, L. Macpherson, J. Mikola, K. Mouhu, A. P. Mähönen, Ü. Niinemets, E. Oksanen, K. Overmyer, E. T. Palva, L. Pazouki, V. Pennanen, T. Puhakainen, P. Poczai, B. J. H. M. Possen, M. Punkkinen, M. M. Rahikainen, M. Rousi, R. Ruonala, C. Van Der Schoot, A. Shapiguzov, M. Sierla, T. P. Sipilä, S. Sutela, T. H. Teeri, A. I. Tervahauta, A. Vaattovaara, J. Vahala, L. Vetchinnikova, A. Welling, M. Wrzaczek, E. Xu, L. G. Paulin, A. H. Schulman, M. Lascoux, V. A. Albert, P. Auvinen, Y. Helariutta, and J. Kangasjärvi. 2017. Genome sequencing and population genomic analyses provide insights into the adaptive landscape of silver birch. Nature Genetics 49:904–912.
Samis, K. E., K. D. Heath, and J. R. Stinchcombe. 2008. Discordant longitudinal clines in flowering time and Phytochrome C in *Arabidopsis thaliana*. Evolution 62:2971–2983.
Samis, K. E., C. J. Murren, O. Bossdorf, K. Donohue, C. B. Fenster, R. L. Malmberg, M.

D. Purugganan, and J. R. Stinchcombe. 2012. Longitudinal trends in climate drive flowering time clines in North American *Arabidopsis thaliana*. Ecology and Evolution 2:1162–1180.
Schluter, D., and D. Nychka. 1994. Exploring fitness landscapes. American Naturalist 143:597–616.
Schneider, A. C., T. D. Lee, M. A. Kreiser, and G. T. Nelson. 2014. Comparative and interactive effects of reduced precipitation frequency and volume on the growth and function of two Perennial grassland species. International Journal of Plant Sciences 175:702–712.
Schneider, A., M. A. Friedl, and D. Potere. 2009. A new map of global urban extent from MODIS satellite data. Environmental Research Letters 4.
Schwinning, S., and J. R. Ehleringer. 2001. Water use trade-offs and optimal adaptations to pulse-driven arid ecosystems. Journal of Ecology 89:464–480.
Serieys, L. E. K., A. Lea, J. P. Pollinger, S. P. D. Riley, and R. K. Wayne. 2015. Disease and freeways drive genetic change in urban bobcat populations. Evolutionary Applications 8:75–92.
Sexton, J. P., S. Y. Strauss, and K. J. Rice. 2011. Gene flow increases fitness at the warm edge of a species' range. Proceedings of the National Academy of Science USA 108:11704–9.
Shafer, A. B. A., and J. B. W. Wolf. 2013. Widespread evidence for incipient ecological speciation: A meta-analysis of isolation-by-ecology. Ecology Letters 16:940–950.
Shaw, R. G., C. J. Geyer, S. Wagenius, H. H. Hangelbroek, and J. R. Etterson. 2008. Unifying life-history analyses for inference of fitness and population growth. American Naturalist 172:E35-47.
Shepherd, J. M., and H. Pierce. 2002. Rainfall modification by major urban areas: Observations from spaceborne rain radar on the TRMM satellite. Journal of Applied Meteorology 41:689–701.
Siepielski, A. M., M. B. Morrissey, M. Buoro, S. M. Carlson, C. M. Caruso, S. M. Clegg, T. Coulson, J. Dibattista, K. M. Gotanda, C. D. Francis, J. Hereford, J. G. Kingsolver, N. Sletvold, E. I. Svensson, M. J. Wade, and A. D. C. Maccoll. 2017. Precipitation drives global variation in natural selection. Science 962:959–962.
Simão, F. A., R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, and E. M. Zdobnov. 2015. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31:3210–3212.
Slabbekoorn, H., and M. Peet. 2003. Birds sing at a higher pitch in urban noise. Nature 424:267.
Slatkin, M. 1987. Gene flow and the geographic structure of natural populations. Science 236:787–792.
Smoliak, B. V., P. K. Snyder, T. E. Twine, P. M. Mykleby, and W. F. Hertel. 2015. Dense network observations of the Twin Cities Canopy-Layer urban heat island. Journal of Applied Meteorology and Climatology 54:1899–1917.
Stach, A., M. Smith, C. A. Skjøth, and J. Brandt. 2007. Examining *Ambrosia* pollen episodes at Poznań (Poland) using back-trajectory analysis. International Journal of Biometeorology 51:275–86.
Stephen, R. K., R. S. Dexter, N. Maurine, M. W. Lorne, and R. T. Douglas. 2009. Adaptation and colonization history affect the evolution of clines in two introduced

species. New Phytologist 9999:678–690.
Stinchcombe, J. R., A. F. Agrawal, P. A. Hohenlohe, S. J. Arnold, and M. W. Blows. 2008. Estimating nonlinear selection gradients suing quadratic regression coefficients: Double or nothing? Evolution 62:2435–2440.
Stinchcombe, J. R., and H. E. Hoekstra. 2008. Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. Heredity 100:158–70.
Stinchcombe, J. R., C. Weinig, M. Ungerer, K. M. Olsen, C. Mays, S. S. Halldorsdottir, M. D. Purugganan, and J. Schmitt. 2004. A latitudinal cline in flowering time in *Arabidopsis thaliana* modulated by the flowering time gene FRIGIDA. Proceedings of the National Academy of Sciences USA 101:4712–7.
Stinson, K. A., J. M. Albertine, L. M. S. Hancock, T. G. Seidler, and C. A. Rogers. 2016. Northern ragweed ecotypes flower earlier and longer in response to elevated CO2: what are you sneezing at? Oecologia 182: 587 - 594
Storey, J. D., and R. Tibshirani. 2003. Proceedings of the National Academy of Sciences USA 100:9440–5.
Supek, F., M. Bošnjak, N. Škunca, and T. Šmuc. 2011. Revigo summarizes and visualizes long lists of gene ontology terms. PLoS One 6.
Team, R. C. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
Thomas, P. D., V. Wood, C. J. Mungall, S. E. Lewis, and J. A. Blake. 2012. On the use of gene ontology annotations to assess functional similarity among orthologs and paralogs: A short report. PLoS Computational Biology 8:1–7.
Thompson, K. A., M. Renaudin, and M. T. J. Johnson. 2016. Urbanization drives parallel adaptive clines in plant populations. Proceedings of the Royal Society of London B B 8:1–51.
Tiffin, P., and J. Ross-Ibarra. 2014. Advances and limits of using population genetics to understand local adaptation. Trends in Ecology and Evolution 29: 673-680
Toole, E. H., and E. Brown. 1946. Final results of the Duvel buried seed experiment. Journal of Agricultural Research 72:201–210.
Travisano, M., and R. G. Shaw. 2013. Lost in the map. Evolution 67:305–14.
Trenberth, K. E., D. Aiguo, R. M. Rasmussen, and D. M. Parsons. 2003. The changing character of precipitation. Bulletin of the American Meteorological Society. 84:1205.
Turesson, G. 1922. The genotypical response of the plant species to the habitat. Hereditas.
van Boheemen, L. A., E. Lombaert, K. A. Nurkowski, B. Gauffre, L. H. Rieseberg, and K. A. Hodgins. 2017. Multiple introductions, admixture and bridgehead invasion characterize the introduction history of *Ambrosia artemisiifolia* in Europe and Australia. Molecular Ecology 26:5421–5434.
Vieira, F. G., M. Fumagalli, A. Albrechtsen, and R. Nielsen. 2013. Estimating inbreeding coefficients from NGS data: Impact on genotype calling and allele frequency estimation. Genome Research 23:1852–1861.
Walck, J. L., S. N. Hidayati, K. W. Dixon, K. Thompson, and P. Poschlod. 2011. Climate change and plant regeneration from seed. Global Change Biology 17:2145–2161.
Wang, I. J., R. E. Glor, and J. B. Losos. 2013. Quantifying the roles of ecology and

geography in spatial genetic divergence. Ecology Letters 16:175–82.
Wang, Z., M. Gerstein, and M. Snyder. 2009. RNA-Seq: a revolutionary tool for transcriptomics. Nature Reviews Genetics 10:57–63.
Way, D. A., and R. A. Montgomery. 2015. Photoperiod constraints on tree phenology, performance and migration in a warming world. Plant, Cell and the Environment 38:1725–1736.
Weber, J. N., B. K. Peterson, and H. E. Hoekstra. 2013. Discrete genetic modules are responsible for complex burrow evolution in *Peromyscus* mice. Nature 493:402–5.
Weltzin, J. J. F., M. E. M. Loik, S. Schwinning, D. G. Williams, P. a. Fay, B. M. Haddad, J. Harte, T. E. Huxman, A. K. Knapp, G. Lin, W. T. Pockman, M. R. Shaw, E. E. Small, M. D. Smith, S. D. Smith, D. T. Tissue, and J. C. Zak. 2003. Assessing the response of terrestrial ecosystems to potential changes in precipitation. Bioscience 53:941.
Whitlow, T. H., N. L. Bassuk, and D. L. Reichert. 1992. A 3-Year study of water relations of urban street trees. Journal of Applied Ecology 29:436–450.
Wilczek, A. M., M. D. Cooper, T. M. Korves, and J. Schmitt. 2014. Lagging adaptation to warming climate in *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences USA 111:7906–13.
Willemsen, R. W. 1975. Effect of stratification temperature and germination temperature on germination and the induction of secondary dormancy in common ragweed seeds. American Journal of Botany 62:1–5.
Williams, J. W., B. N. Shuman, T. Webb III, P. J. Bartlein, and P. L. Leduc. 2004. Late quaternary vegetation dynamics in North America: scaling from taxa to biomes. Ecological Monographs 74:309–334.
Winchell, K. M., R. G. Reynolds, S. R. Prado-Irwin, A. R. Puente-Rol, and L. J. Revell. 2016. Phenotypic shifts in urban areas in the tropical lizard Anolis cristatellus. Evolution70:1009–1022.
Wolfe, M. D., and S. J. Tonsor. 2013. Adaptation to spring heat and drought in northeastern Spanish *Arabidopsis thaliana*. New Phytologist 201: 323 - 334.
Wright, S. 1943. Isolation by distance. Genetics 114: 114-138.
Wu, Z., P. Dijkstra, G. W. Koch, J. Peñuelas, and B. A. Hungate. 2011. Responses of terrestrial ecosystems to temperature and precipitation change: A meta-analysis of experimental manipulation. Global Change Biology 17:927–942.
Wuebbles, D. J., D. W. Fahey, K. A. Hibbard, D. J. Dokken, B. C. Stewart, and T. K. Maycock. 2017. Climate Science Special Report: Fourth Natiaonl Climate Assessment. Washington, DC, USA.
Xu, H., S. Qiang, Z. Han, J. Guo, Z. Huang, H. Sun, S. He, H. Ding, H. Wu, and F. Wan. 2006. The status and causes of alien species invasion in China. Biodiversity and Conservation 15:2893–2904.
Yakub, M., and P. Tiffin. 2016. Living in the city: Urban environments shape the evolution of a native annual plant. Global Change Biology 23: 2082-2089
Yamada, Y. 1962. Genotype by environment interaction and genetic correlation of the same trait under different environments. Journal of Japanese Genetics 37:498–509.
Yoder, J. B., J. Stanton-Geddes, P. Zhou, R. Briskine, N. D. Young, and P. Tiffin. 2014. Genomic signature of adaptation to climate in *Medicago truncatula*. Genetics 196:1263–1275.

Yoshida, T., J. Mogami, and K. Yamaguchi-Shinozaki. 2014. Omics approaches toward defining the comprehensive abscisic acid signaling network in plants. Plant Cell and Physiology 56:1043–1052.

Ziska, L. H., D. E. Gebhard, D. a. Frenz, S. Faulkner, B. D. Singer, and J. G. Straka. 2003. Cities as harbingers of climate change: Common ragweed, urbanization, and public health. Journal of Allergy and Clinical Immunology 111:290–295.

Ziska, L., K. Knowlton, C. Rogers, D. Dalan, N. Tierney, M. A. Elder, W. Filley, J. Shropshire, L. B. Ford, C. Hedberg, P. Fleetwood, K. T. Hovanky, T. Kavanaugh, G. Fulford, R. F. Vrtis, J. A. Patz, J. Portnoy, F. Coates, and D. Frenz. 2011. Recent warming by latitude associated with increased length of ragweed pollen season in central North America. Proceedings of the National Academy of Sciences USA 108:4248–4251.

Appendix 1

A. Supplementary text – Model equations

A1. Linear mixed model used to investigate differences among urban and rural seed sources across all field sites. Terms in italics are random effects:

trait = seed weight + site + source region + source region: site + (1|*maternal plant*) + ε

A2. Generalized linear mixed model used to investigate differences among urban and rural seed sources at each field site. Terms in italics are random effects:

trait = seed weight + source region + (1| *maternal plant*) + ε

A3. Linear mixed model used to investigate differences among urban populations and among rural populations across all field sites. Terms in italics are random effects:

trait = seed weight + source population + site + source population: site + (1| source population: *maternal line*) + ε

A4. Selection differential equation used to investigate differences in net selection between urban genotypes at each field site.

w = trait + source region + trait: source region + ε

A5. Selection gradient equation used to investigate differences in direct selection between urban and rural genotypes at each field site:

w = trait1 + trait2 + trait3 + source region + trait1: source region + trait2: source region + trait3: source region + ε

B. Supplementary figures

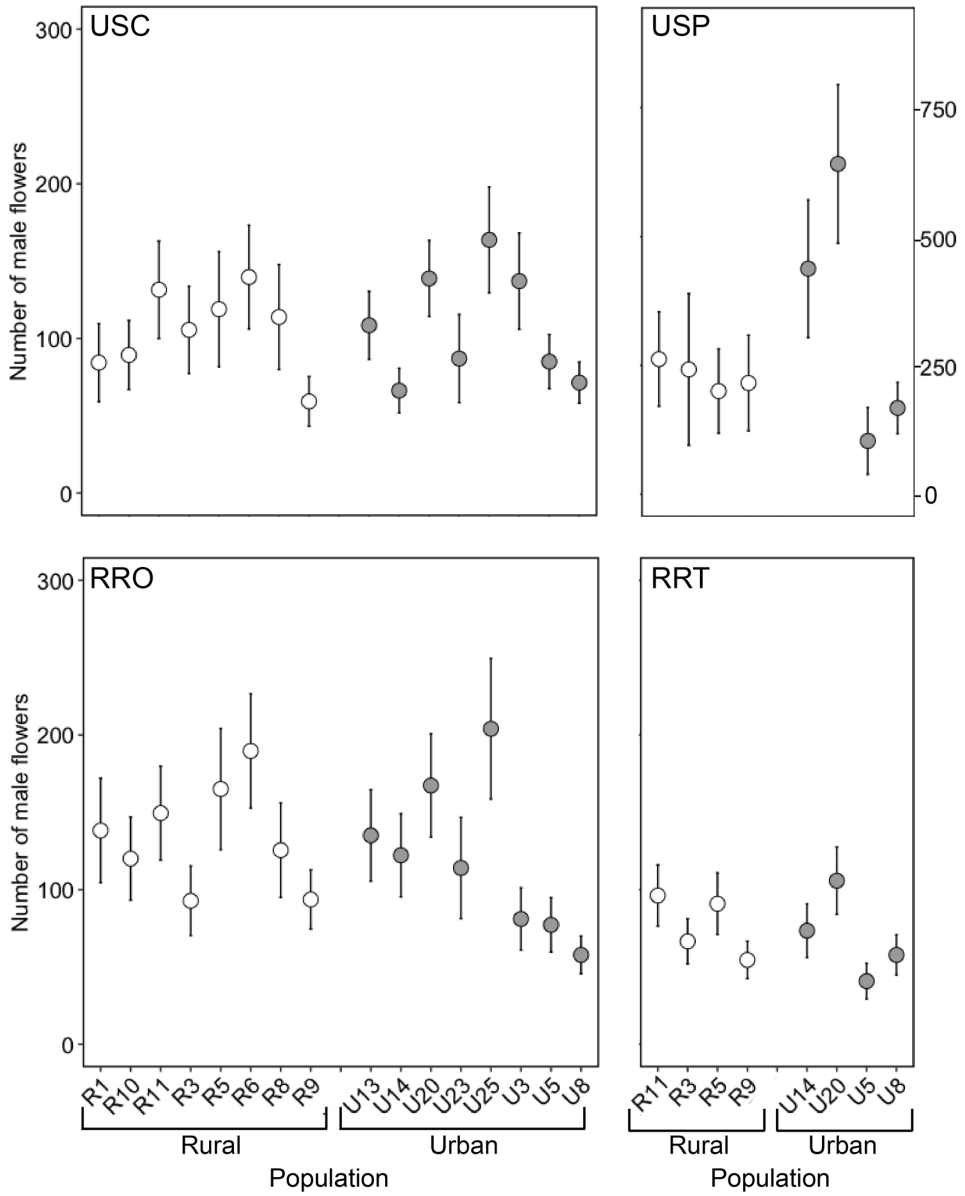

Figure S1. Mean number of male flowers produced by urban and rural populations at each field site. White = rural genotypes, grey = urban genotypes. The bars indicate +/- standard error. Population names are listed on the x-axis. There was a significant effect of population source on urban populations but not rural populations (see Table S7,S8). Means were extracted from generalized linear mixed models. Rural field sites = RRT, RRO; urban field sites = USC, USP. Note the different scale for USP.

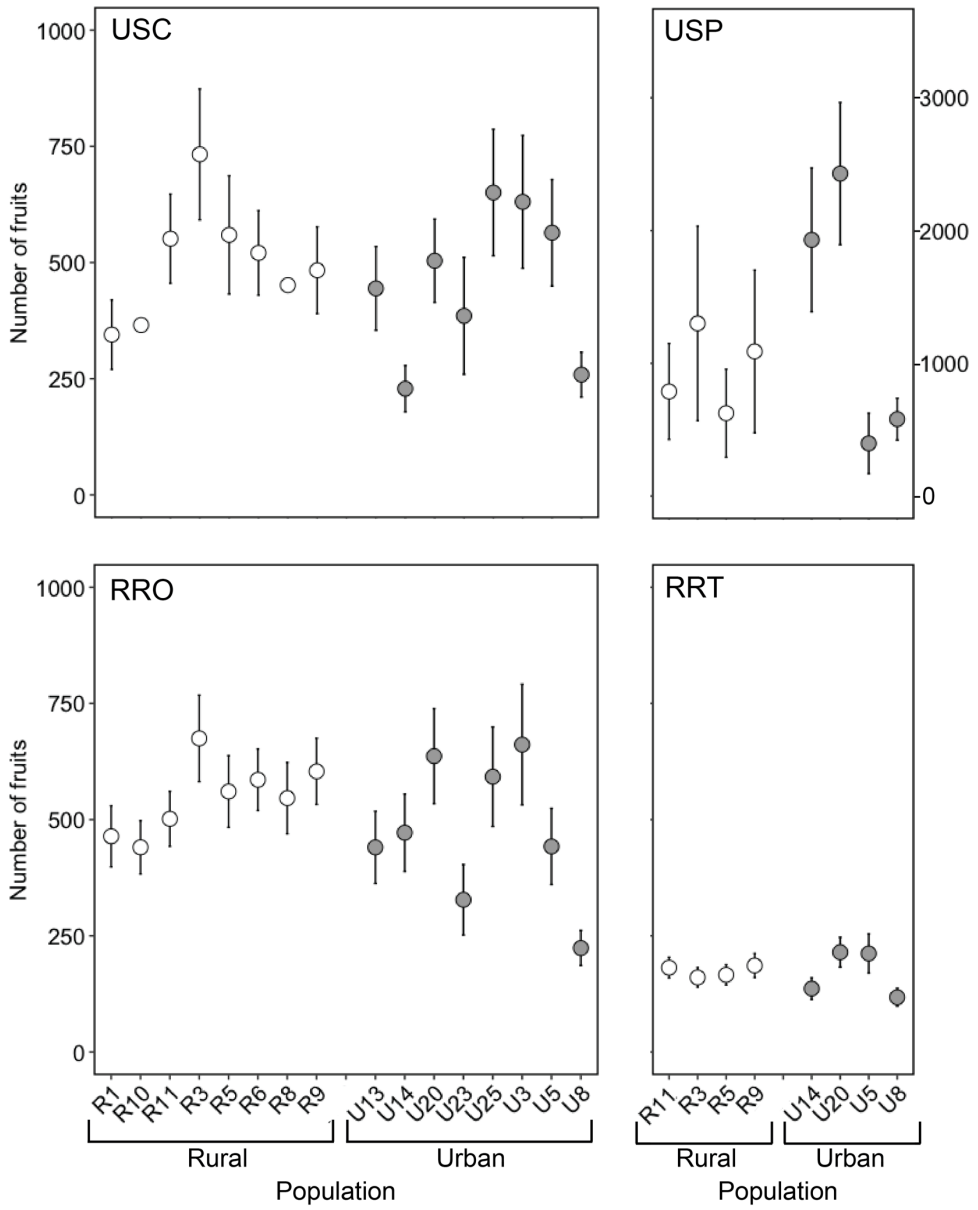

Figure S2. Mean number of fruits produced by urban and rural populations at each field site. White = rural genotypes, grey = urban genotypes. The bars indicate +/- standard error. Population names are listed on the x-axis. There was a significant effect of population source on urban populations but not rural populations (see Table S7, S8). Means were extracted from generalized linear mixed models. Rural field sites = RRT, RRO; urban field sites = USC, USP. Note the different scale for USP.

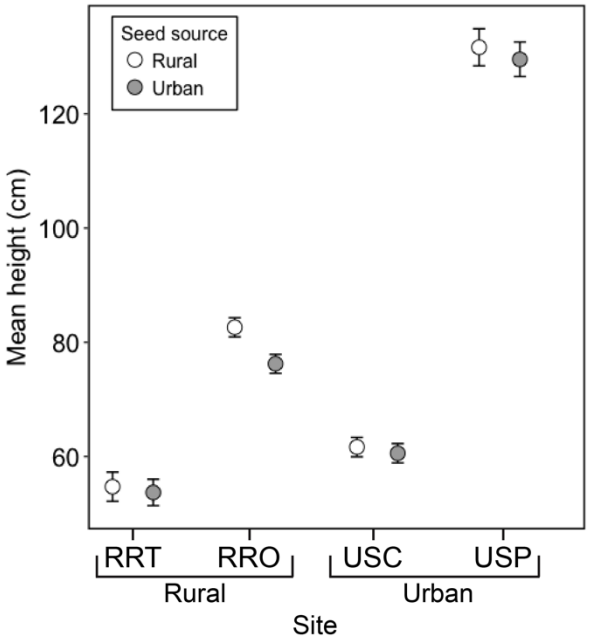

Figure S3. Mean height of urban and rural genotypes at each field site. White = rural genotypes, grey = urban genotypes. The bars indicate +/- standard error. There was a significant effect of source region on height.(see Table S4). Rural field sites = RRT, RRO; urban field sites = USC, USP. Means were extracted from generalized linear mixed models.

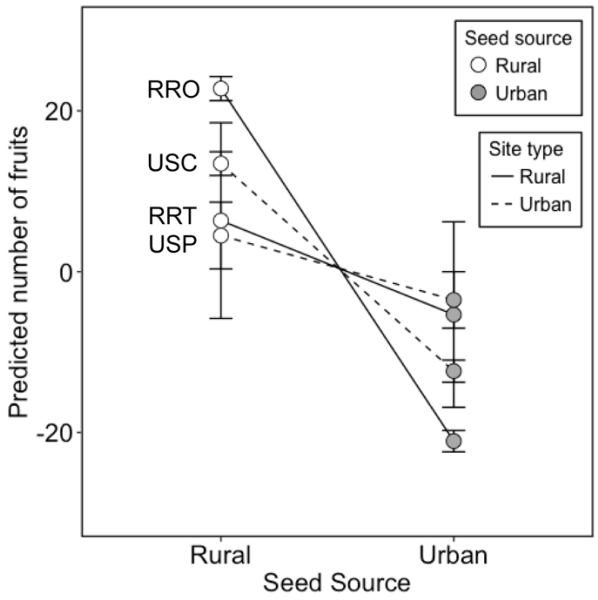

Figure S4. Mean number of predicted fruits produced by urban and rural genotypes at each field site, expressed as deviation from site mean. White = rural genotypes, grey = urban genotypes. Dashed lines = urban field sites, solid lines = rural field sites. The bars indicate +/- standard error on original means. Predicted means were extracted from separate aster models for each site. There was a significant site × source region interaction ($p < 0.0001$, Table S9). Rural field sites = RRT, RRO; urban field sites = USC, USP.

C. Supplementary tables

Table S1. Description of populations used in reciprocal transplant experiment. *Region* indicates whether the population was collected in the urban or rural region. *Habitat* is a description of habitat of the population. *n* indicates the number of maternal plants that were included in each field site. *Garden* indicates which populations were included in the large common gardens (USC, RRO) and which were included in the small gardens (USP, RRT).

Population	Region	Latitude	Longitude	Habitat	Soil	n	Gardens
R1	Rural	44.59418	-93.08842	Roadside	Clay	11	USC, RRO
R3	Rural	44.60452	-93.07011	Roadside	Clay	13	USC, RRO, USP, RRT
R5	Rural	44.58693	-93.11606	Weedy patch	Clay loam	12	USC, RRO, USP, RRT
R6	Rural	44.54348	-93.00144	Roadside	Sandy clay loam	12	USC, RRO
R8	Rural	44.55733	-93.01525	Roadside	Clay	12	USC, RRO
R9	Rural	44.53828	-92.97973	Roadside	Clay	12	USC, RRO, USP, RRT
R10	Rural	44.53361	-93.01993	Roadside	Sandy clay loam	12	USC, RRO
R11	Rural	44.54356	-93.03395	Roadside	Sandy clay loam	12	USC, RRO, USP, RRT
U3	Urban	44.95025	-93.29013	Bike path	Clay	11	USC, RRO
U5	Urban	44.94208	-93.2778	Lawn	Clay	14	USC, RRO, USP, RRT
U8	Urban	45.027	-93.26003	Lawn	Clay loam	12	USC, RRO, USP, RRT
U13	Urban	45.01841	-93.24388	Lawn	Clay loam	12	USC, RRO
U14	Urban	44.96826	-93.26376	Parking lot	-	12	USC, RRO, USP, RRT
U20	Urban	44.95323	-93.21041	Riverbank	Sandy loam	13	USC, RRO, USP, RRT
U23	Urban	44.97101	-93.24968	Lawn	-	8	USC, RRO
U25	Urban	44.96805	-93.22025	Railroad	Sandy loam	12	USC, RRO

Table S2. Land use characterization of populations used in reciprocal transplant experiment. We extracted the land use around each population locale (based on latitude and longitude) from the National Landcover Database (2011). Each column represents the percentage of the indicated land use type in a 100-meter radius. Descriptions of each land use type are included below the table.

Population	Developed, open space	Developed, low intensity	Developed, medium intensity	Developed, high intensity	Pasture/Hay	Cultivated crops	Grassland	Deciduous forest	Open water
R1	0.14	0	0	0	0.03	0.83	0	0	0
R3	0.15	0	0	0	0	0.85	0	0	0
R5	0.11	0.08	0.06	0	0.06	0.65	0	0	0
R6	0.16	0	0	0	0	0.83	0	0	0
R8	0.20	0.03	0	0	0.26	0.51	0	0	0
R9	0.21	0	0	0	0	0.76	0	0.03	0
R10	0.17	0	0	0	0.11	0.50	0.22	0	0
R11	0.20	0	0	0	0	0.31	0.48	0	0
U3	0	0	0.51	0.47	0	0	0	0	0
U5	0.03	0.06	0.59	0.31	0	0	0	0	0
U8	0.30	0.17	0.33	0.19	0	0	0	0	0
U13	0.06	0.40	0.54	0	0	0	0	0	0
U14	0.03	0.06	0.40	0.51	0	0	0	0	0
U20	0.35	0.16	0.11	0.08	0	0	0	0.03	0.27
U23	0.19	0.11	0.38	0.32	0	0	0	0	0
U25	0.06	0.26	0.49	0.20	0	0	0	0	0

Developed, open space – Mostly vegetation (parks, golf course, recreational areas). Impervious surface <20%,
Developed, low intensity – Mix of constructed material and vegetation. Impervious surface 20-49% of cover
Developed, medium intensity - Mix of constructed material and vegetation. Impervious surface 50-79% of cover
Developed, high intensity – Highly developed areas (apartments, row houses, industrial). Impervious surface 80-100% of cover
Pasture/hay – areas of grasses or legume planted for grazing or productions of seeds or hay crops.
Cultivated crops – areas used for production of annul crops such as corn, soybeans, vegetables, cotton, orchards
Deciduous forest – areas dominated by trees generally taller than 5 m tall and greater than 20% of total vegetation cover
Open water – all areas of open water, with less than 25% cover of vegetation or soil

Table S3. Likelihood ratio test results for effect of source region (urban vs rural) and site on phenology traits. Each term was evaluated by comparing nested models with and without the term of interest. Models were run using mixed models in lme4. Degrees of freedom of the likelihood ratio test = 1.

Trait	Factor	Change in LL	χ^2	P-value
Transition to reproduction	Site	2.4	4.79	0.19
	Region	1.3	2.58	0.11
	Site × Region	0.4	0.85	0.84
First open male flower	Site	35.8	71.54	**< 0.0001**
	Region	2.2	4.29	**0.04**
	Site × Region	1.5	3.00	0.39
First open female flower	Site	4.8	209.6	**< 0.0001**
	Region	5.8	11.71	**0.0006**
	Site × Region	0.7	1.40	0.70
Male flower to female flower	Site	10	20.04	**0.0001**
	Region	0.8	1.45	0.23
	Site × Region	0.9	1.88	0.60

Terms in bold are significant at $p < 0.05$

Change LL = change in log likelihood between the simple (term excluded) and complex model (term included)

χ^2 = chi-square statistic

Table S4. Likelihood ratio test results for effect of source region (urban vs rural) on fitness and size. Analyses were parsed by field site and models were run using generalized linear mixed models in lme4, with a poisson error distribution and a log link function for fitness, and a Gaussian distribution and identity link function for height. Degrees of freedom of the likelihood ratio test = 1.

	Trait	Change in LL	χ^2	P-value
Urban sites				
USC	Male flowers	0	0.003	0.95
	Fruits	1	1.01	0.31
	Height	0.1	0.20	0.65
USP	Male flowers	0.8	1.48	0.22
	Fruits	1	1.53	0.22
	Height	0.1	0.25	0.62
Rural sites				
RRO	Male flowers	0	1.41	0.24
	Fruits	2	2.53	0.11
	Height	3.7	7.44	**0.006**
RRT	Male flowers	0.1	0.19	0.66
	Fruits	0.1	0.30	0.59
	Height	0.05	0.09	0.76

Terms in bold are significant at $p < 0.05$

Change LL = change in log likelihood between the simple (term excluded) and complex model (term included)

χ^2 = chi-square statistic

Table S5. Local-foreign contrasts for male and female fitness (GLMMs). Contrasts were calculated separately for each site from least-square means extracted from generalized linear mixed models in lme4. Local source and foreign source indicate which genotype was the local and foreign genotype at each field site, respectively.

	Trait	Local source	Foreign source	Contrast
Urban sites				
USC	Male flowers	Urban	Rural	0.77
	Fruits	Urban	Rural	-52.87
USP	Male flowers	Urban	Rural	102.33
	Fruits	Urban	Rural	440.79
Rural sites				
RRO	Male flowers	Rural	Urban	18.59
	Fruits	Rural	Urban	72.31
RRT	Male flowers	Rural	Urban	5.81
	Fruits	Rural	Urban	11.47

Table S6. Likelihood ratio test results for effect of population source and site on phenology of urban and rural populations. Data was subset separately into urban and rural populations. Each term was evaluated by comparing nested models with and without the term of interest. Models were run using mixed models in lme4. Degrees of freedom of the likelihood ratio test = 1.

	Factor	Change in LL	χ^2	P value
Urban populations				
Transition to reproduction	Site	0.5	0.91	0.82
	Population	23.3	46.52	**< 0.0001**
	Site × Population	2.9	5.87	0.95
First open male flower	Site	21.9	43.65	**< 0.0001**
	Population	19	37.85	**< 0.0001**
	Site × Population	5.3	10.77	0.63
First open female flower	Site	51.7	103.34	**< 0.0001**
	Population	21	41.93	**< 0.0001**
	Site × Population	8	15.93	0.25
Male flower to female flower	Site	3.7	7.42	**0.060**
	Population	20.1	20.12	**0.005**
	Site × Population	6.4	12.73	0.47
Rural populations				
Transition to reproduction	Site	2.7	5.39	0.14
	Population	8.5	17.06	**0.02**
	Site × Population	4.1	8.22	0.82
First open male flower	Site	16.3	32.61	**< 0.0001**
	Population	6.6	13.12	0.07
	Site × Population	3.7	7.53	0.87
First open female flower	Site	55.3	110.71	**< 0.0001**
	Population	12.4	24.82	**0.0008**
	Site × Population	6.6	13.09	0.44
Male flower to female flower	Site	12.6	12.58	**0.006**
	Population	8.2	16.32	**0.02**
	Site × Population	10.8	21.66	0.06

Terms in bold are significant at $p < 0.05$

Change LL = change in log likelihood between the simple (term excluded) and complex model (term included)

χ^2 = chi-square statistic

Table S7. Likelihood ratio test results for effect of urban population source on fitness and height of urban populations. Data was subset into urban populations and parsed by site. Analyses were run using generalized linear mixed models in lme4, with a poisson error distribution and a log link function. Degrees of freedom of the likelihood ratio test = 1.

	Trait	Change in LL	χ^2	P-value
Urban sites				
USC	Male flowers	8.3	16.57	**0.02**
	Fruits	12	22.48	**0.002**
	Height	28	54.56	**< 0.0001**
USP	Male flowers	6.4	12.87	**0.005**
	Fruits	8	16.29	**0.001**
	Height	4.16	8.32	**0.04**
Rural sites				
RRO	Male flowers	10.7	21.48	**0.003**
	Fruits	11	22.91	**0.002**
	Height	21	42.12	**< 0.0001**
RRT	Male flowers	7.8	7.80	**0.05**
	Fruits	4.6	9.17	**0.03**
	Height	18.6	37.12	**< 0.0001**

Terms in bold are significant at $p < 0.05$

Change LL = change in log likelihood between the simple (term excluded) and complex model (term included)

χ^2 = chi-square statistic

Table S8. Likelihood ratio test results for effect of rural population source on fitness and height of rural populations. Data was subset into rural populations and parsed by site. Analyses were run using generalized linear mixed models in lme4, with a poisson error distribution and a log link function. Degrees of freedom of the likelihood ratio test = 1.

	Trait	Change in LL	χ^2	P-value
Urban sites				
USC	Male flowers	3.9	7.76	0.35
	Fruits	5	9.72	0.20
	Height	5.55	11.10	0.13
USP	Male flowers	0.5	0.85	0.81
	Fruits	0	1.07	0.78
	Height	4.24	8.47	**0.04**
Rural sites				
RRO	Male flowers	4.8	9.45	0.22
	Fruits	4	7.66	0.36
	Height	3.93	7.86	0.34
RRT	Male flowers	2.1	4.34	0.22
	Fruits	0.4	0.84	0.84
	Height	1.74	3.48	0.32

Terms in bold are significant at $p < 0.05$.
Change LL = change in log likelihood between the simple (term excluded) and complex model (term included)
χ^2 = chi-square statistic

Table S9. Summary of aster model comparisons testing the effect of site, source region, and their interaction on individual male and female lifetime fitness. The terminal fitness node for male and female fitness were set as number of fruits and number of flowers, respectively. Analyses were run using fixed effect aster models. Significance of each term was determined by comparing sequentially nested models with likelihood ratio tests. Degrees of freedom of the likelihood ratio test = 1.

	Residual df	Deviance	P
Female fitness			
Region	3	5.4	**0.02**
Site	3	188322	**< 0.0001**
Region x Site	7	517.51	**<0.0001**
Male fitness			
Region	3	0.43	0.5094
Site	3	31276	**<0.0001**
Region x Site	7	241.7	**<0.0001**

Terms in bold are significant at $p < 0.05$.

Table S10. Local-foreign contrasts for male and female fitness (aster models). Contrasts were calculated separately for each site from predicted female and male lifetime fitness, estimated from aster models. Local source and foreign source indicate which genotype was the local and foreign genotype at each field site, respectively.

	Trait	Local source	Foreign source	Contrast
Urban sites				
USC	Male fitness	Urban	Rural	-4.19
	Female fitness	Urban	Rural	-25.78
USP	Male fitness	Urban	Rural	26.34
	Female fitness	Urban	Rural	-8.00
Rural sites				
RRO	Male fitness	Rural	Urban	13.26
	Female fitness	Rural	Urban	43.84
RRT	Male fitness	Rural	Urban	12.26
	Female fitness	Rural	Urban	9.69

Table S11. Genotypic correlations among traits at urban field sites. USC is above the diagonal and USP is below the diagonal. Correlations in bold are significant at $p < 0.05$.

	Transition to reproduction	**First open male flower**	**First open female flower**	**Male flower to female flower**	**Height**
Transition to reproduction		**0.66**	**0.41**	**-0.53**	-0.11
First open male flower	**0.59**		**0.70**	**-0.69**	-0.10
First open female flower	**0.41**	**0.78**		0.03	0.04
Male flower to female flower	**-0.40**	**-0.56**	0.07		0.14
Height	-0.02	-0.002	0.08	0.15	

Table S12. Genotypic correlations among traits at rural field sites. RRO is above the diagonal and RRT is below the diagonal.

	Transition to reproduction	**First open male flower**	**First open female flower**	**Male flower to female flower**	**Height**
Transition to reproduction		**0.69**	**0.52**	**-0.42**	**0.17**
First open male flower	**0.66**		**0.68**	**-0.63**	0.08
First open female flower	**0.41**	**0.80**		0.11	**0.22**
Male flower to female flower	**-0.45**	**-0.51**	0.08		0.09
Height	-0.19	-0.08	-0.07	0.06	

Table S13. Quadratic selection differentials from linear regression models.

		Transition to reproduction		First open male flower		Male flower to female flower		Height	
Urban		Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural
USC	S_{male}	**-0.44**	-0.29	-0.13	**-0.34**	0.17	**0.29**	**0.34**	**0.51**
	C_{male}	**-0.22**	-0.06	-0.04	-0.04	-0.01	-0.01	0.01	-0.05
	S_{female}	**-0.37**	-0.21	-0.05	**-0.23**	-0.01	0.11	**0.26**	**0.45**
	C_{female}	**-0.21**[a]	0.06[a]	-0.03	0.013	0.05	**0.05**	-0.04	-0.09
USP	S_{male}	0.07	0.32	-0.01	-0.20	0.18	0.63	0.33*	0.14
	C_{male}	-0.16	0.13	-0.19	-0.11	0.07	-0.13	-0.07	-0.08
	S_{female}	0.07	0.39	-0.01	-0.23	0.13	-0.03	0.23	0.08
	C_{female}	0.15	0.16	-0.16	-0.12	0.17	-0.04	-0.10	-0.12
Rural									
RRO	S_{male}	**-0.36**	-0.07	**-0.29**	-0.10	**0.44**	**0.24**	**0.44**	**0.28**
	C_{male}	**-0.29**[a]	0.11[a]	-0.14	0.08	-0.15	-0.02	0.06	-0.09
	S_{female}	**-0.32**	0.04	**-0.29**	0.00	**0.18**	0.02	**0.33**	0.08
	C_{female}	0.05	0.14	**-0.16**	**0.11**	0.04	**0.08**	0.04	0.08
RRT	S_{male}	-0.08	-0.22	-0.20	0.10	0.12	0.02	**0.42**	0.27
	C_{male}	0.05	-0.06	0.03	0.08	0.03	0.16	-0.10	0.04
	S_{female}	-0.27	-0.27	**-0.34**	-0.07	-0.16	0.07	**0.31**	**0.32**
	C_{female}	-0.03	0.01	0.07	0.05	0.25	0.09	-0.14	0.17

S = linear regression coefficient, C = quadratic regression coefficient.

Selection differentials were calculated separately on male and female fitness

Selection differentials in bold are significant at $p < 0.05$.

[a] Selection differentials which are significantly different between urban and rural seed sources at $p<0.05$

Table S14. Quadratic selection gradients from linear regression models.

		Transition to reproduction		Time to first open male flower		Male flower to female flower	
Urban sites		Urban	Rural	Urban	Rural	Urban	Rural
USC	$S\text{-}B_{male}$	**-0.43**	-0.22	-0.08	**-0.25**	0.08	0.15
	$C\text{-}B_{male}$	**-0.44**	-0.12	0.10	0.04	-0.08	0.00
	$S\text{-}B_{female}$	**-0.46**	-0.25	-0.07	**-0.22**	-0.14	-0.01
	$C\text{-}B_{female}$	**-0.44**	-0.26	-0.10	-0.02	0.14	**0.16**
USP	$S\text{-}B_{male}$	0.32	0.44	0.18	-0.15	0.25	0.03
	$C\text{-}B_{male}$	-0.24	0.60	-0.64	-0.18	0.30	-0.40
	$S\text{-}B_{female}$	0.36	0.47	0.6	-0.22	0.16	-0.18
	$C\text{-}B_{female}$	-0.24	0.72	-0.96	-0.20	0.60	-0.26
Rural sites							
RRO	$S\text{-}B_{male}$	-0.17	0.06	-0.02	0.08	**0.40**	**0.24**
	$C\text{-}B_{male}$	-0.14	0.24	**-0.34**[a]	0.18[a]	-0.26	-0.12
	$S\text{-}B_{female}$	**-0.26**	0.06	-0.03	0.10	0.22	0.05
	$C\text{-}B_{female}$	-0.20	0.06	-0.30	0.08	-0.10	0.12
RRT	$S\text{-}B_{male}$	-0.13	-0.30	-0.22[a]	0.28[a]	-0.02	-0.06
	$C\text{-}B_{male}$	0.10	-0.46	0.32	**0.52**	0.12	-0.02
	$S\text{-}B_{female}$	**-0.51**	0.50	**-0.52**[a]	0.06[a]	**-0.47**	-0.14
	$C\text{-}B_{female}$	0.06	-0.26	-0.06	0.42	0.38	-0.10

B = multiple regression coefficients. S = linear multiple regression coefficient, C = quadratic multiple regression coefficient

Selection gradients were calculated separately on male and female fitness

Selection gradients in bold are significant at $p < 0.05$.

[a] Selection gradients which are significantly different between urban and rural seed sources at $p<0.05$

Note: Quadratic selection gradients are multiplied by two, as per Stinchcombe et al. 2008.

Appendix 2

A. Supplementary figures

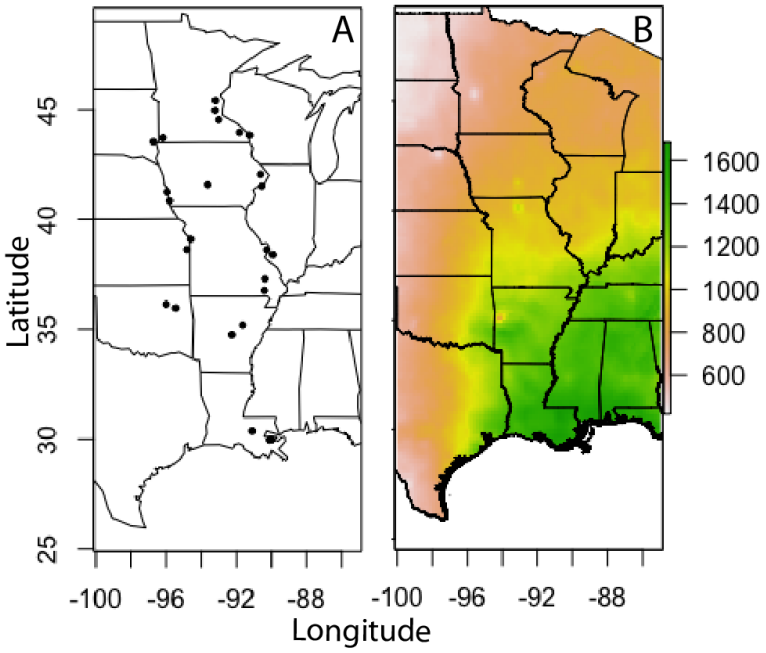

Figure S1. Map of A) population sampling locations included in the experiment, and B) annual precipitation. Each black dot in panel A represents a separate population. The legend for annual precipitation in B) corresponds to mm, with annual precipitation increasing from northern to southern latitudes.

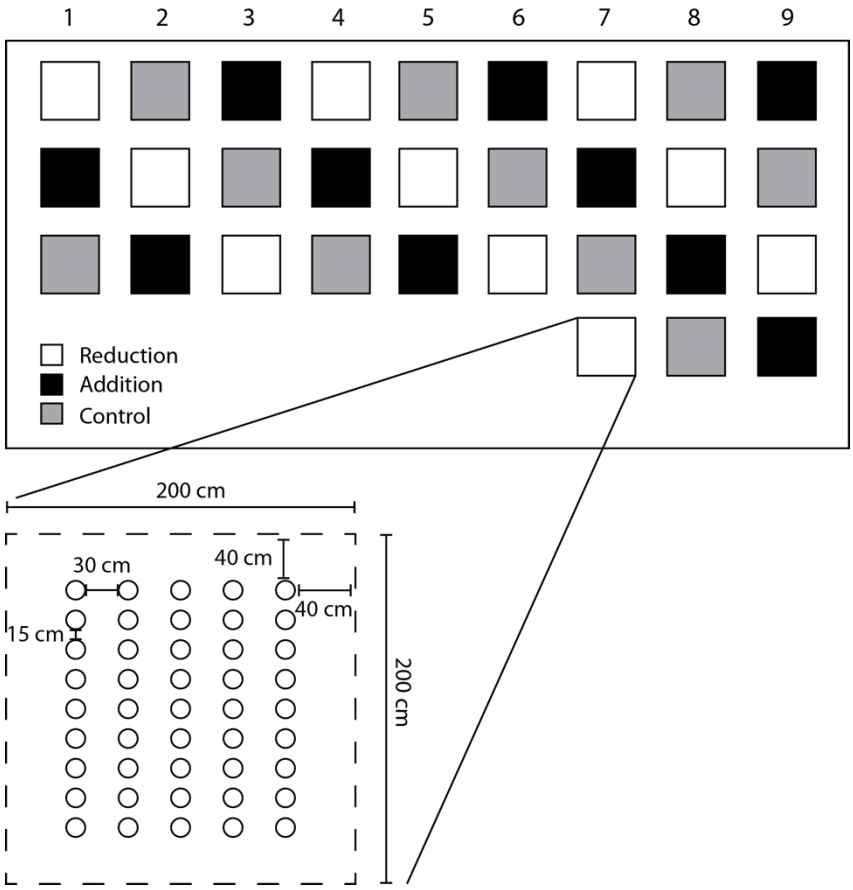

Figure S2. Layout of precipitation treatments and plots. Each square represents a different plot. Plots were spaced with 1 m on every side. The numbers across the top of the diagram indicate block number or the column number in which each plot occurred. Inset displays individual plot layout, where a circle represents an individual plant.

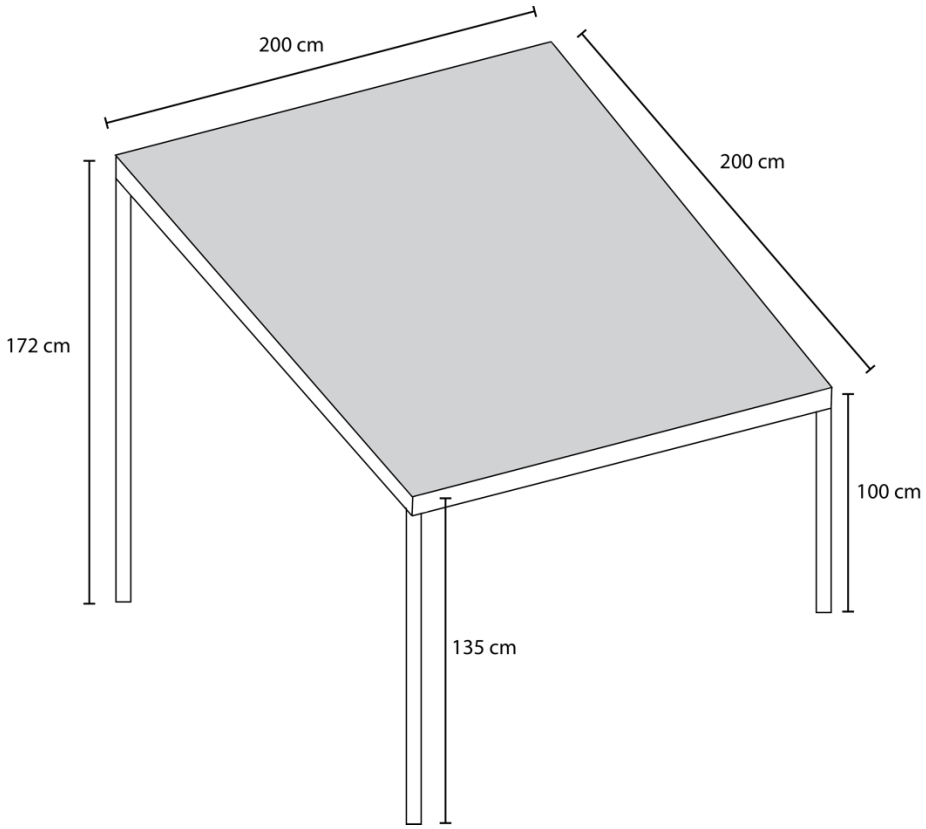

Figure S3. Rainout shelter design and materials. All material shown in white was constructed out of 1'' white plastic PVC pipes with assorted fittings. The roof was constructed out of greenhouse over-wintering plastic (shown in grey, AT Super 1 UV clear 3 mil film, Greenhouse Megastore) and was stretched over the PVC frame and kept in place with Snap Clamps (Greenhouse Megastore). When we wanted to exclude a rainfall event, we attached the roof to the PVC legs via t-joints.

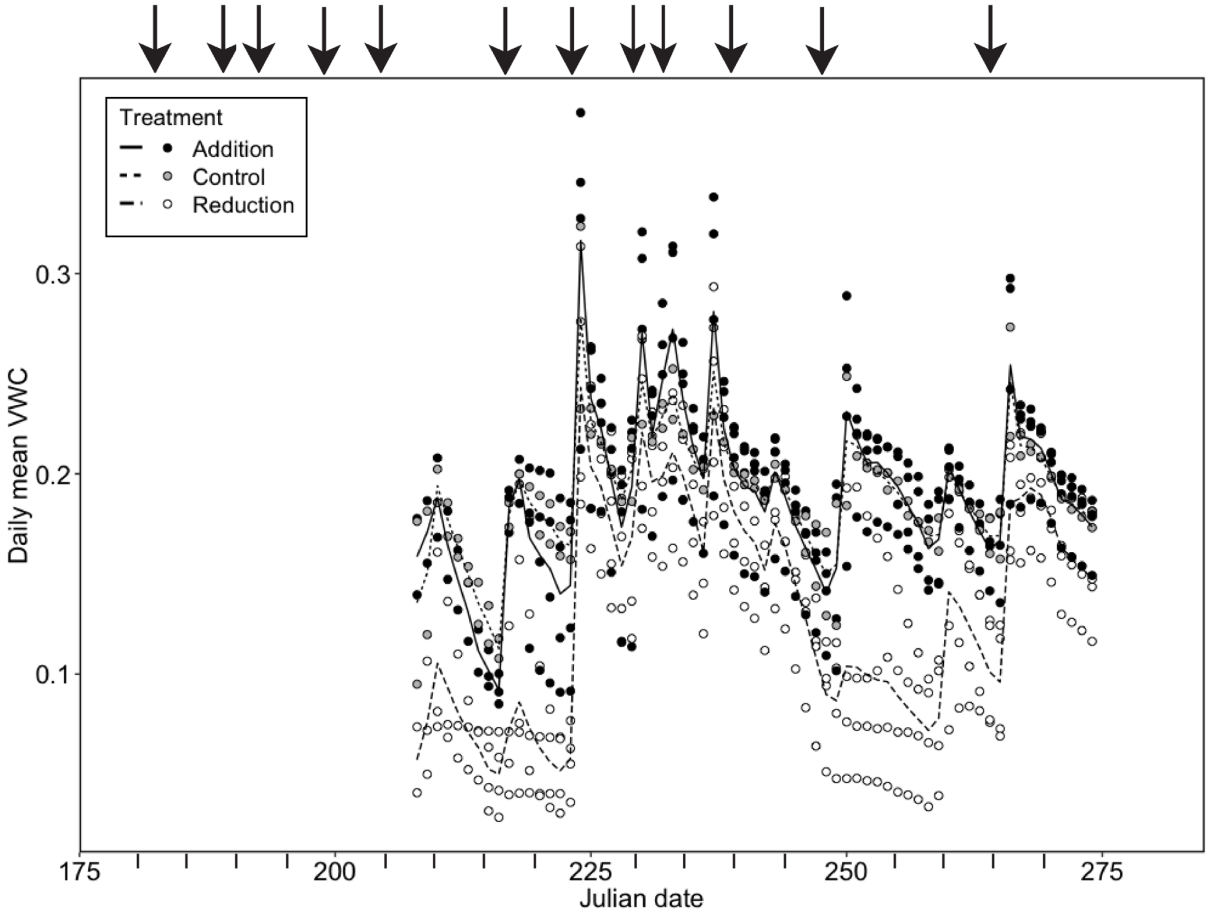

Figure S4. Volumetric water content (VWC) of soil during August and September 2016. VWC was collected hourly for 10 plots (4 reduction, 4 addition, 2 control) using ECH_20-E5 Decagon soil moisture sensors and Em50 ECH_20 data loggers. Each data point is the VWC average for each probe on a given day. The lines correspond to the average VWC across all probes for a given treatment. The black arrows at the top of the figure indicate when rainfall events were excluded and when additions occurred. Data were not available for the first month of the experiment due to data logger malfunction.

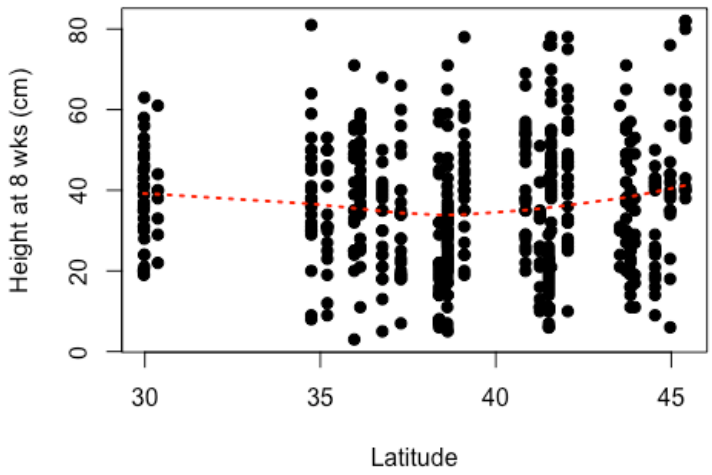

Figure S5. Individual height at 8 weeks versus latitude. The red dotted line is a spline fit to the relationship between the two variables using *smooth.spline*().

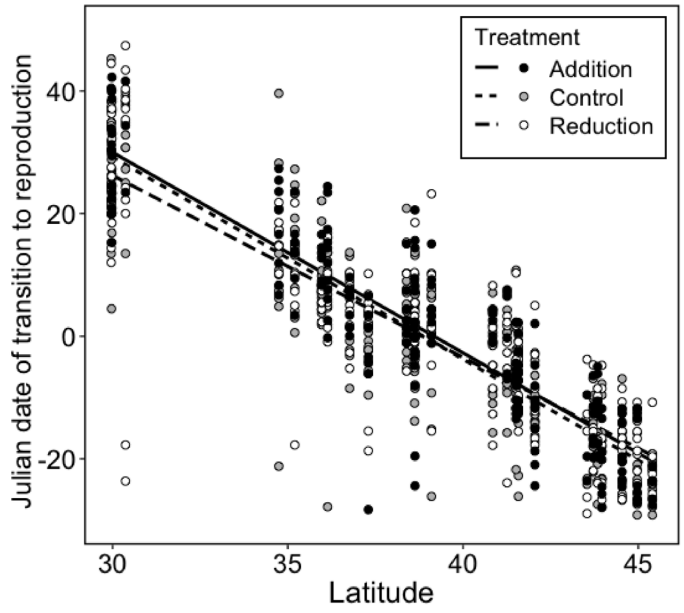

Figure S6. Phenology was strongly associated with latitude, with populations from northern latitudes initiating reproduction earlier than those from southern latitudes. Residuals were extracted from linear models with only block and edge as predictors. Linear regression lines are shown for Julian date of transition to reproduction vs latitude for each rainfall treatment.

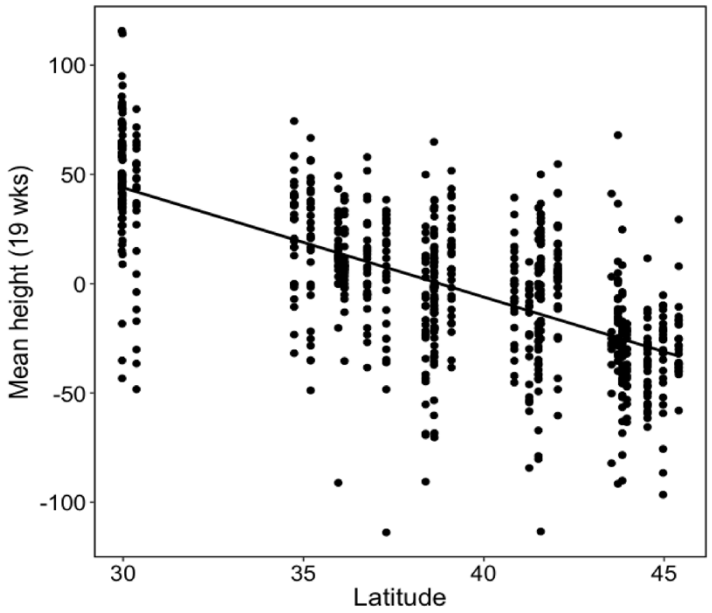

Figure S7. Populations from southern latitudes tended to be taller. Residuals were extracted from linear models with only block and edge as predictors. Linear regression line is shown for height at 19 weeks vs latitude.

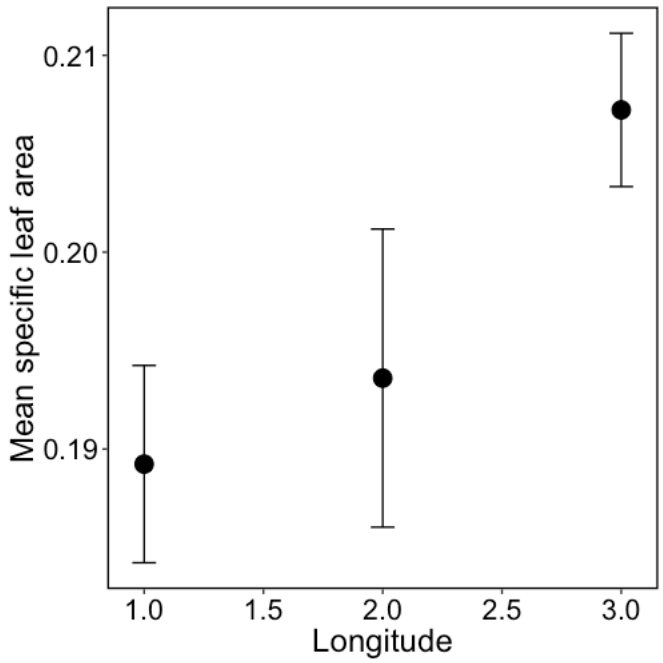

Figure S8. Western populations had significantly lower SLA than eastern populations. 1 = longitudes > 94, 2 = longitudes < 94 > 92, 3 = longitudes < 92. Least-square means were extracted from a linear model and are presented with standard error bars.

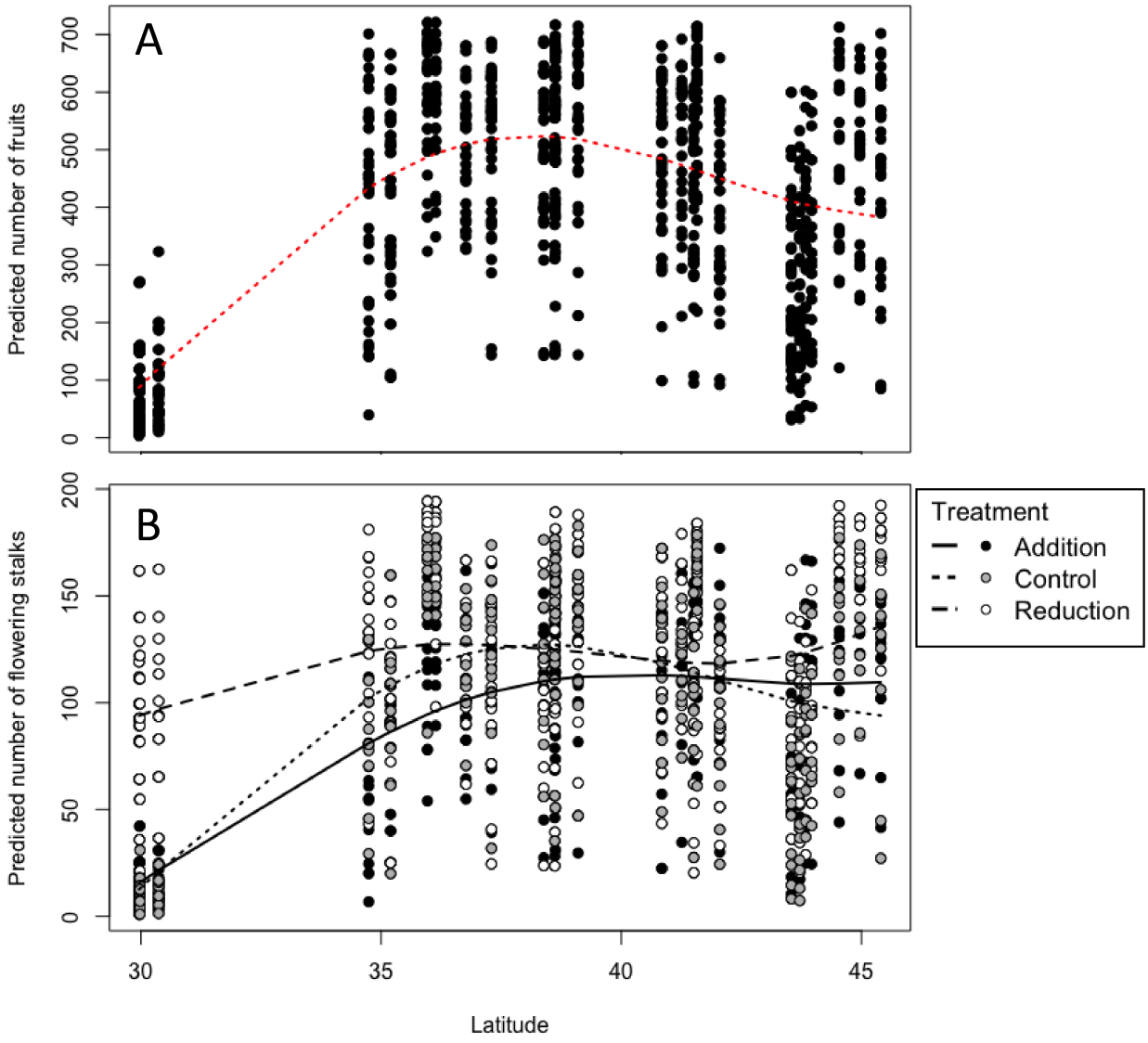

Figure S9. Individual data for latitude vs predicted fitness values from *aster* for a) female fitness, b) male fitness. d) height at 8 weeks. The red dotted line is a spline fit to the relationship between the two variables using *smooth.spline*().

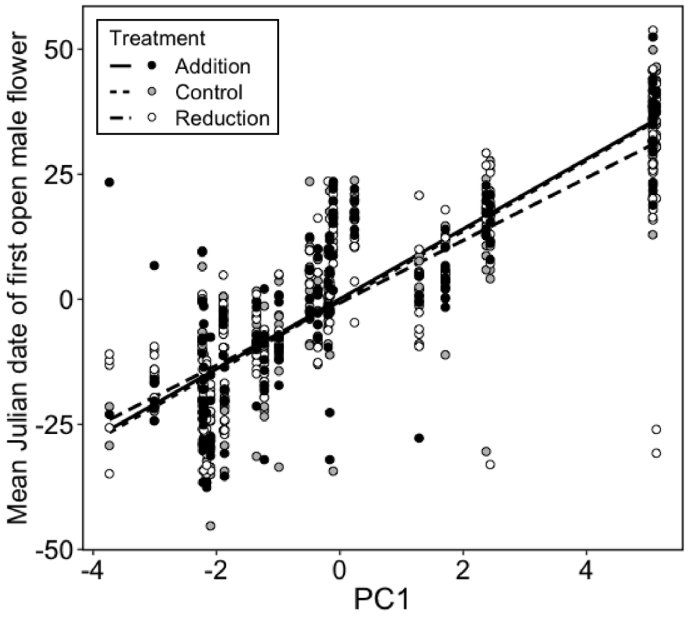

Figure S10. Phenology was strongly associated with PC1, with populations from wetter climates flowering earlier in response to the rainfall reduction treatment than populations from drier climates. Residuals were extracted from linear models with only block and edge as predictors. Linear regression lines are shown for Julian date of first open male flower vs PC1 for each rainfall treatment.

B. Supplementary tables

Table S1. Description of populations used in precipitation manipulation experiment. Populations are listed in ascending order by latitude. *n* indicates the number of maternal lines that were included from that population.

Population	Latitude	Longitude	*n*
LA4	29.9626	-90.1095	24
LA6	29.97096	-90.01688	16
LA1	29.9815	-90.02892	18
BBN	30.36979	-91.10567	24
LTR	34.74136566	-92.25947634	23
LDV	35.19796198	-91.64370698	25
WAG	35.96535395	-95.45782392	25
ARK	36.13697554	-95.99441688	19
POP	36.76623098	-90.40547492	15
CLW	37.29567794	-90.36880106	25
SR9	38.38645	-89.92686	25
SU8	38.61736	-90.26415	25
PAO	38.62834563	-94.81757954	25
KAC	39.10268604	-94.60068525	25
MPL	40.84571607	-95.80212162	25
OMA	41.25688838	-95.94715796	12
ROI	41.51026297	-90.56311849	23
DMS	41.57788835	-93.62835764	25
PRC	42.05130841	-90.63232642	25
SFA	43.53882148	-96.72653811	24
BMO	43.71449138	-96.18077593	25
LAC	43.83766655	-91.2513499	25
LEW	43.96234593	-91.83442563	18
R10	44.53656	-93.02021	25
U25	44.96804	-93.22023	18
R16	45.40264	-93.19718	25

Table S2. Variance explained by the first four PCs. The PCA included all BIOCLIM precipitation variables (BIO12 - BIO19). Standard deviation is also shown.

	PC1	**PC2**	**PC3**	**PC4**
Variance	0.847	0.118	0.026	0.006
St. dev	2.60	0.970	0.455	0.220

Table S3. BIOCLIM loadings on first four PCs. BIO12 = annual precipitation, BIO13 = precipitation of the wettest month, BIO14 = precipitation of the driest month, BIO15 = precipitation seasonality, BIO16 = precipitation of the wettest quarter, BIO17 = precipitation of the driest quarter, BIO18 = precipitation of the warmest quarter, BIO19 = precipitation of the coldest quarter.

	PC1	**PC2**	**PC3**	**PC4**
BIO12	0.384			
BIO13	0.348	0.312	0.597	0.418
BIO14	0.362	-0.325		-0.407
BIO15	-0.337	0.455	0.220	-0.666
BIO16	0.361	0.285	0.358	-0.303
BIO17	0.375	-0.214	-0.132	-0.115
BIO18	0.269	0.670	-0.645	
BIO19	0.378	-0.118	-0.176	-0.314

Table S4. Results of linear model testing the effects of latitude, longitude, and treatment on phenology and size traits. Non-linearity was also tested by including latitude^2. F-values and statistical significance are based on type II sums of squares. Terms in bold are significant at p = 0.05. The asterisks indicate the level of significance after performing a sequential Bonferroni to account for multiple testing.

Trait	Factor	SS	df	F-value	P-value
Transition to reproduction	Block	1467	8	2.97	**0.002****
	Edge	28	1	0.46	0.50
	Treatment	156	2	1.26	0.28
	Latitude	388	1	6.27	**0.01***
	Longitude	388	2	3.14	**0.04**
	Latitude^2	7	1	0.11	0.73
	Latitude × Longitude	357	2	2.89	0.06
	Treatment × Latitude	635	2	5.14	**0.006***
	Treatment × Longitude	120	4	0.48	0.75
	Residuals	43247	700		
First open male flower	Block	948	8	1.62	0.11
	Edge	50	1	0.68	0.41
	Treatment	311	2	2.13	0.12
	Latitude	487	1	6.66	**0.01***
	Longitude	1792	2	12.26	**<0.0001*****
	Latitude^2	3	1	0.04	0.84
	Latitude × Longitude	61	2	0.41	0.66
	Treatment × Latitude	574	2	0.41	**0.02**
	Treatment × Longitude	116	4	0.40	0.81
	Residuals	48445	663		
First open female flower	Block	1051	8	1.99	**0.05**
	Edge	5020	1	0.30	0.58
	Treatment	89	2	0.68	0.51
	Latitude	244	1	3.69	**0.05**
	Longitude	1398	2	10.58	**<0.0001*****
	Latitude^2	13	1	0.20	0.65
	Latitude × Longitude	67	2	0.51	0.60
	Treatment × Latitude	160	2	1.21	0.30
	Treatment × Longitude	211	4	0.80	0.52
	Residuals	45598	690		
Height (8 wks)	Block	16382	8	9.53	**<0.0001*****
	Edge	379	1	1.76	0.18
	Treatment	3638	2	8.46	**0.0002*****
	Latitude	2412	1	11.22	**0.0008****
	Longitude	10800	2	25.12	**<0.0001*****
	Latitude^2	965	1	4349	**0.03**
	Latitude × Longitude	71	2	0.16	0.85
	Treatment × Latitude	505	2	1.17	0.31
	Treatment × Longitude	907	4	1.05	0.38
	Residuals	153037	712		
Height (19 wks)	Block	38325	8	6.72	**<0.0001*****
	Edge	2587	1	3.63	0.06
	Treatment	7	2	0.00	0.99
	Latitude	2949	1	4.14	**0.04**
	Longitude	5018	2	3.52	**0.03**
	Latitude^2	54	1	0.07	0.78
	Latitude × Longitude	3095	2	2.17	0.11
	Treatment × Latitude	255	2	0.18	0.84
	Treatment × Longitude	516	4	0.12	0.95
	Residuals	502965	706		

Specific leaf area	Block	0.250	8	6.88	**<0.0001*** **
	Edge	0.006	1	1.41	0.23
	Treatment	0.038	2	4.15	**0.02**
	Latitude	0.000	1	0.00	0.94
	Longitude	0.048	2	5.32	**0.005***
	Latitude^2	0.009	1	1.93	0.16
	Latitude × Longitude	0.022	2	2.53	0.08
	Treatment × Latitude	0.000	2	0.01	0.99
	Treatment × Longitude	0.016	4	0.90	0.46
	Residuals	2.960	652		

*** < 0.001

**< 0.01

* < 0.05

Table S5. Results of linear model testing the effect of treatment on phenology and size traits. F-values and statistical significance are based on type II sums of squares. The asterisks indicate the level of significance after performing a sequential Bonferroni to account for multiple testing.

Trait	Factor	SS	df	F-value	P-value
Transition to reproduction	Block	4609	8	1.97	**0.05**
	Edge	11	1	0.04	0.84
	Treatment	125	2	0.21	0.81
	Residual	209132	715		
First open male flower	Block	3092	8	0.92	0.50
	Edge	39	1	0.09	0.76
	Treatment	202	2	0.24	0.78
	Residual	283973	678		
First open female flower	Block	3304	8	1.13	0.34
	Edge	2	1	0.01	0.94
	Treatment	271	2	0.37	0.69
	Residual	257852	705		
Height (8 wks)	Block	15401	8	8.38	**<0.0001*****
	Edge	407	1	1.77	0.18
	Treatment	3246	2	7.07	**0.0009****
	Residual	167148	728		
Height (19 wks)	Block	41632	8	4.00	**0.0001*****
	Edge	1818	1	1.40	0.24
	Treatment	186	2	0.07	0.93
	Residual	940094	722		
Specific leaf area	Block	0.25	8	6.63	**<0.0001*****
	Edge	0.01	1	1.65	0.29
	Treatment	0.05	2	5.35	**0.005***
	Residual	3.16	668		

*** < 0.001

**< 0.01

* < 0.05

Table S6. Summary of aster models testing the effect of treatment on male and female lifetime fitness. The terminal fitness nodes for male and female fitness were set as number of fruits and number of flowering spikes, respectively. Analyses were run using fixed effect aster models. Significance of each term was determined by comparing sequentially nested models with likelihood ratio tests (df = 1).

Factor	Change in deviance	P-value
Female fitness		
Edge	55.6	**< 0.0001**
Block	79.3	**< 0.0001**
Treatment	6.5	**0.04**
Male fitness		
Edge	21.8	**<0.0001**
Block	101.0	**<0.0001**
Treatment	21.7	**<0.0001**

Table S7. Results of linear model testing the effects of latitude, longitude, and treatment on phenology and size traits. F-values and statistical significance are based on type II sums of squares. The asterisks indicate the level of significance after performing a sequential Bonferroni to account for multiple testing.

Trait	Factor	SS	df	F-value	P-value
Transition to reproduction	Block	1522	8	3.07	**0.002****
	Edge	22	1	0.36	0.55
	Treatment	153	2	1.24	0.29
	Latitude	135526	1	2190	**<0.0001*****
	Longitude	545	2	4.40	**0.01***
	Latitude × Longitude	309	2	2.50	0.08
	Treatment × Latitude	647	2	5.23	**0.005***
	Treatment × Longitude	106	4	0.43	0.79
	Residuals	43380	701		
First open male flower	Block	928	8	1.59	0.12
	Edge	46	1	0.63	0.42
	Treatment	333	2	2.28	0.10
	Latitude	191679	1	2622	**<0.0001*****
	Longitude	2343	2	16.03	**<0.0001*****
	Latitude × Longitude	94	2	0.64	0.53
	Treatment × Latitude	608	2	4.16	**0.02**
	Treatment × Longitude	126	4	0.43	0.79
	Residuals	48528	664		
First open female flower	Block	1046	8	1.98	**0.05**
	Edge	18	1	0.27	0.60
	Treatment	104	2	0.79	0.45
	Latitude	172807	1	2612	**<0.0001*****
	Longitude	1950	2	14.7	**<0.0001*****
	Latitude × Longitude	180	2	1.36	0.26
	Treatment × Latitude	176	2	1.33	0.26
	Treatment × Longitude	213	4	0.81	0.52
	Residuals	45713	691		
Height (8 wks)	Block	16067	8	9.42	**<0.0001*****
	Edge	472	1	2.21	0.14
	Treatment	3745	2	8.78	**0.0002*****
	Latitude	1943	1	9.11	**0.002****
	Longitude	11239	2	26.4	**<0.0001*****
	Latitude × Longitude	1399	2	3.28	**0.04**
	Treatment × Latitude	487	2	1.14	0.32
	Treatment × Longitude	989	4	1.16	0.33
	Residuals	152032	713		
Height (19 wks)	Block	37721	8	6.66	**<0.0001*****
	Edge	2776	1	3.92	**0.05**
	Treatment	6	2	0.00	0.99
	Latitude	376993	1	532	**<0.0001*****
	Longitude	5149	2	3.63	**0.03***
	Latitude × Longitude	3779	2	2.67	0.07
	Treatment × Latitude	224	2	0.16	0.85
	Treatment × Longitude	590	4	0.21	0.93
	Residuals	500617	707		
Specific leaf area	Block	0.248	8	6.82	**<0.0001*****
	Edge	0.007	1	1.50	0.22
	Treatment	0.040	2	4.17	**0.01***
	Latitude	0.016	1	3.53	0.07

	Longitude	0.047	2	5.20	**0.006***
	Latitude × Longitude	0.014	2	1.58	0.21
	Treatment × Latitude	0.000	2	0.01	0.99
	Treatment × Longitude	0.017	4	0.94	0.44
	Residuals	2.967	653		

*** < 0.001

**< 0.01

* < 0.05

Table S8. Summary of aster model comparisons testing the effect of latitude, longitude, treatment and their interaction on individual male and female lifetime fitness. Non-linearity was also tested by including latitude^2.The terminal fitness node for male and female fitness were set as number of fruits and number of flowering spikes, respectively. Analyses were run using fixed effect aster models. Significance of each term was determined by comparing sequentially nested models with likelihood ratio tests (df = 1).

Factor	**Change in deviance**	**P-value**
Female fitness		
Edge	58.3	**< 0.0001**
Block	85.4	**< 0.0001**
Latitude	25.6	**<0.0001**
Latitude^2	160.2	**<0.0001**
Longitude	23.6	**< 0.0001**
Treatment	6.7	**0.03**
Latitude × Treatment	1.9	0.38
Longitude × Treatment	10.6	**0.03**
Latitude × Longitude	33.2	**<0.0001**
Male fitness		
Edge	22.4	**<0.0001**
Block	105.47	**<0.0001**
Latitude	18.0	**<0.0001**
Latitude^2	50.4	**<0.0001**
Longitude	21.3	**<0.0001**
Treatment	22.5	**<0.0001**
Latitude × Treatment	15.6	**0.0004**
Longitude × Treatment	3.1	0.54
Latitude × Longitude	39.2	**<0.0001**
Latitude^2 × Treatment	13.8	**0.001**

Table S9. Results of linear model testing the effects of PC1 (all BIOCLIM precipitation variables) on phenology and size traits. F-values and statistical significance are based on type II sums of squares. The asterisks indicate the level of significance after performing a sequential Bonferroni to account for multiple testing.

Trait	Factor	SS	df	F-value	P-value
Transition to reproduction	Block	1601	8	2.24	**0.02**
	Edge	42	1	0.46	0.49
	Treatment	179	2	1.00	0.37
	PC1	143927	1	1611	**<0.0001*****
	PC1 × Treatment	650	2	3.63	**0.03**
	Residual	63329	709		
First open male flower	Block	982	8	1.00	0.43
	Edge	107	1	0.88	0.35
	Treatment	501	2	2.04	0.13
	PC1	199356	1	1627	**<0.001*****
	PC1 × Treatment	872	2	3.56	**0.03**
	Residual	82322	672		
First open female flower	Block	900	8	1.03	0.41
	Edge	32	1	0.29	0.59
	Treatment	187	2	0.85	0.43
	PC1	179863	1	1644	**<0.0001*****
	PC1 × Treatment	285	2	1.30	0.27
	Residual	76451	699		
Height (8 wks)	Block	15498	8	8.41	**<0.0001*****
	Edge	465	1	2.02	0.15
	Treatment	3057	2	6.64	**0.001****
	PC1	30	1	0.13	0.72
	PC1 × Treatment	191	2	0.41	0.66
	Residual	166008	721		
Height (19 wks)	Block	36987	8	6.09	**<0.0001*****
	Edge	2767	1	3.65	0.06
	Treatment	4	2	0.00	0.99
	PC1	386797	1	510	**<0.0001*****
	PC1 × Treatment	238	2	0.16	0.85
	Residual	542299	715		
Specific leaf area	Block	0.262	8	7.14	**<0.0001*****
	Edge	0.008	1	1.71	0.19
	Treatment	0.040	2	4.41	**0.01***
	PC1	0.000	1	0.21	0.65
	PC1 × Treatment	0.001	2	0.14	0.87
	Residual	3.048	661		

*** < 0.001

**< 0.01

* < 0.05

Table S10. Summary of aster model comparisons testing the effect of the precipitation PC1, treatment and their interaction on male and female lifetime fitness. The terminal fitness node for male and female fitness were set as number of fruits and number of flowers, respectively. Analyses were run using fixed effect aster models. Significance of each term was determined by comparing sequentially nested models with likelihood ratio tests. Degrees of freedom of the likelihood ratio test = 1. Terms in bold are significant at $p < 0.05$.

Factor	Change in deviance	P-value
Female fitness		
Edge	57.8	**< 0.0001**
Block	84.2	**< 0.0001**
Treatment	6.8	**0.03**
PC1	71.8	**<0.0001**
PC1 × Treatment	1.6	0.44
Male fitness		**<0.0001**
Edge	22.3	**<0.0001**
Block	104.7	**<0.0001**
Treatment	22.4	**<0.0001**
PC1	52.3	**<0.0001**
PC1 × Treatment	7.7	**0.02**

Appendix 3

A. Supplementary figures

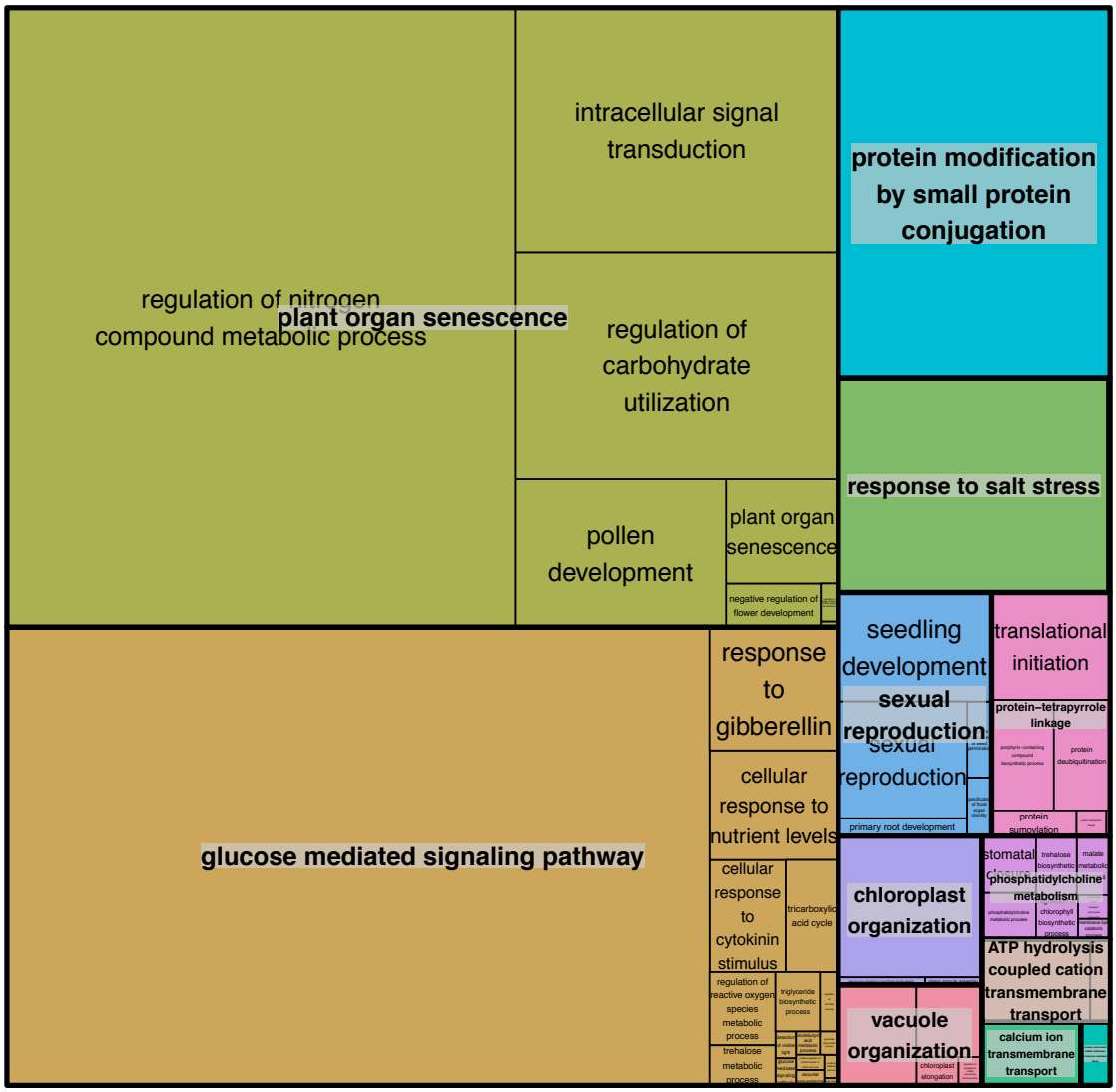

Figure S1. Summary of the biological processes for 176 GO terms enriched for latitude. Figure adapted from REVIGO. Each rectangle represents a cluster of related GO terms labeled according to a representative term. Rectangles are grouped into 'superclusters' of loosely related terms (identified with the same colour) based on SimRel semantic similarity measure.

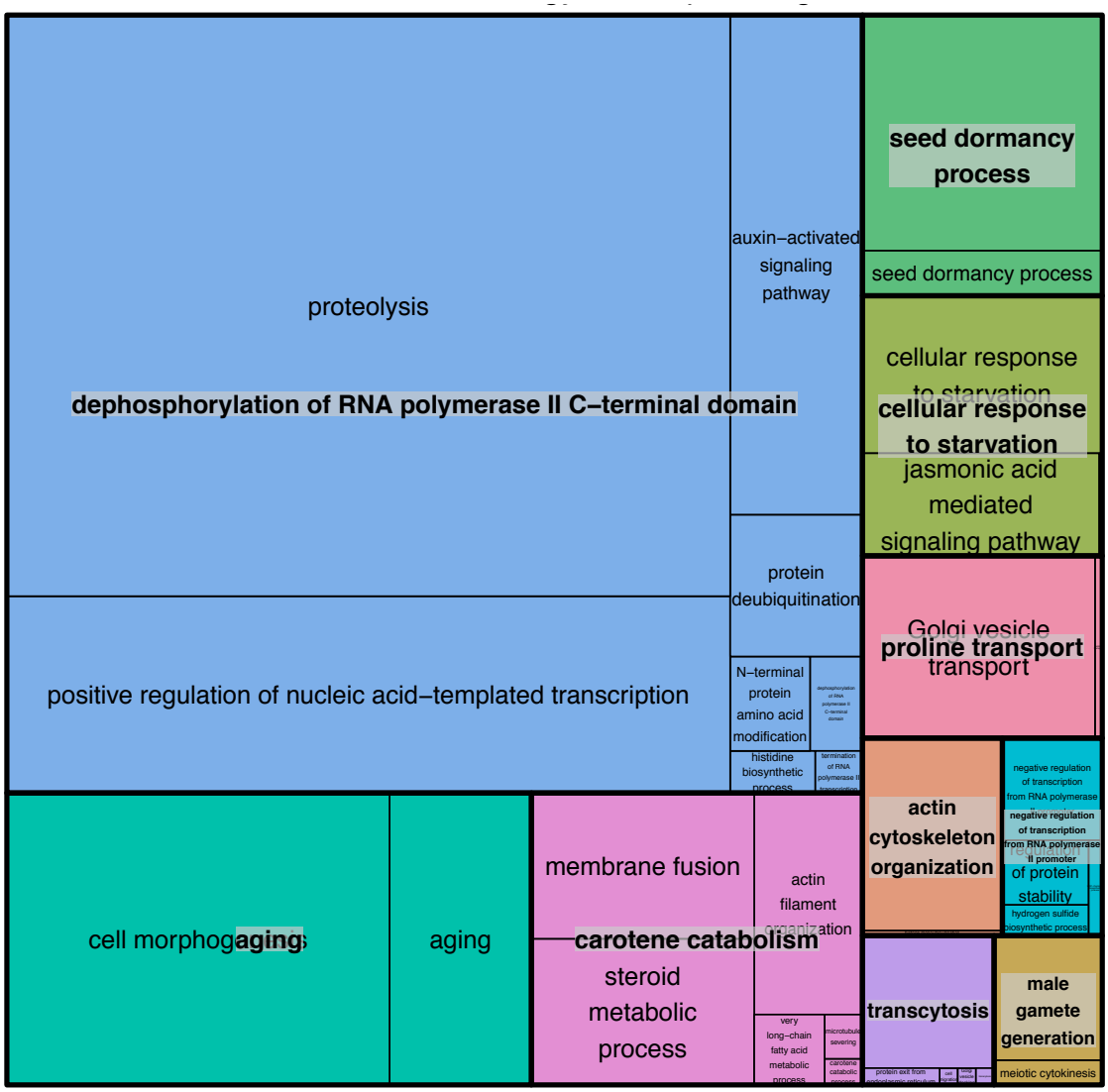

Figure S2. Summary of the biological processes for 101 GO terms enriched for longitude. Figure adapted from REVIGO. Each rectangle represents a cluster of related GO terms labeled according to a representative term. Rectangles are grouped into 'superclusters' of loosely related terms (identified with the same colour) based on SimRel semantic similarity measure.

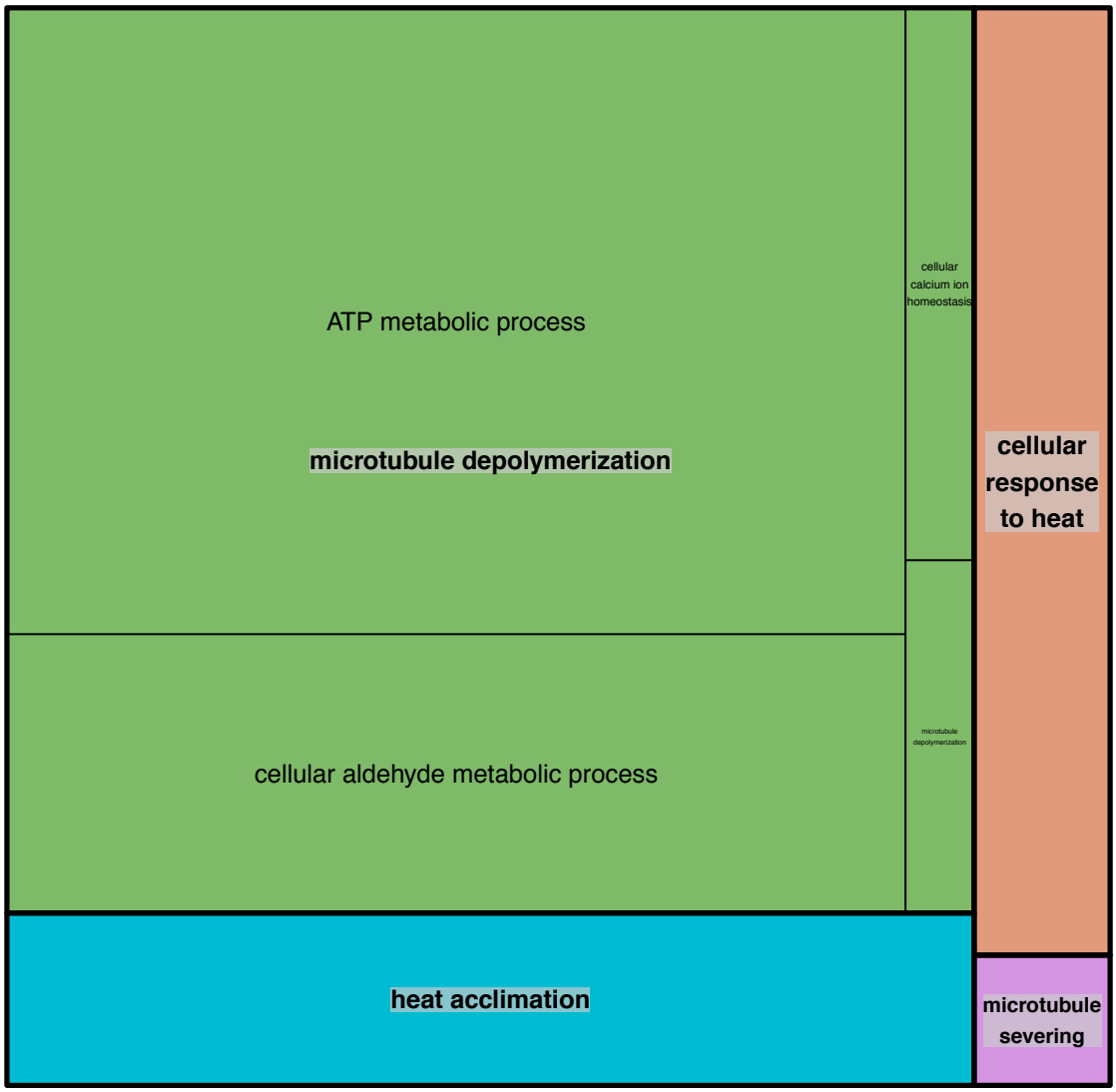

Figure S3. Summary of the biological processes for 23 GO terms enriched for female flowering time after controlling for latitude. Figure adapted from REVIGO. Each rectangle represents a cluster of related GO terms labeled according to a representative term. Rectangles are grouped into 'superclusters' of loosely related terms (identified with the same colour) based on SimRel semantic similarity measure.

B. Supplementary tables

Table A1. *Arabidopsis thaliana* homologs and associated functions of the top *A. artemisiifolia* contigs containing SNP outliers from LFMM2 analysis of latitude, longitude, and female flowering time. The top 50 unique contigs containing outlier SNPs with q-values below 0.05 were examined for latitude and longitude, and all outliers associated with flowering time. Each contig sequence was BLASTED against *A. thaliana* and the homolog with the highest e-value was extracted. Any contigs that did not have matches or had e-values below e-05 are not included. Functions and descriptions were inferred from TAIR and association publications. * Indicates genes discussed in the text.

Variable	Contig	Q-value	*A. thaliana* locus	Gene name	E-value	Description and function
Latitude	TRINITY_DN16529_c1_g1_i2	1.49E-10	AT4G33950	ATOST1*	1.19E-115	stomatal opening/closing
Latitude	TRINITY_DN15287_c2_g1_i1	3.73E-09	AT2G46790	APRR9*	2.94E-09	circadian clock; components of temperature-sensitive circadian system
Latitude	TRINITY_DN15392_c0_g1_i2	1.15E-08	AT2G21470	ATSAE2	1.87E-44	protein sumolation; embryo development ending in seed dormancy
Latitude	TRINITY_DN12859_c0_g1_i1	4.62E-08	AT3G02720	DJ1D	4.59E-10	encodes a glyoxalase
Latitude	TRINITY_DN17383_c2_g2_i1	6.75E-08	AT3G14400	UBP25	2.72E-06	encodes a ubiquitin-specific protease
Latitude	TRINITY_DN3511_c0_g1_i1	8.06E-08	AT4G12520		7.78E-05	bifunctional inhibitor; lipid-transfer protein; seed storage 2S albumin superfamily protein
Latitude	TRINITY_DN15767_c0_g1_i3	8.06E-08	AT5G58330	NADP-MDH	1.25E-78	lactate/malate dehydrogenase family protein
Latitude	TRINITY_DN17792_c1_g1_i7	1.34E-07	AT4G36920	AP2*	1.22E-26	floral homeotic gene; involved in specification of floral organ and meristem identity; development of ovule and seed coat; controls seed mass
Latitude	TRINITY_DN15479_c1_g1_i1	1.50E-07	AT1G19110		3.10E-06	inter-alpha-trypsin heavy chain-like protein
Latitude	TRINITY_DN18637_c1_g1_i4	1.99E-06	AT1G17220	FUG1	3.62E-62	functions as translation initiation factor; involved in leaf variegation

Latitude	TRINITY_DN18058_c1_g2_i2	3.17E-06	AT4G14540	NF-YB3*	5.01E-43	positive regulation of transcription; nuclear factor Y; water-use efficiency + drought tolerance (Han et al. 2013)
Latitude	TRINITY_DN16074_c0_g1_i1	4.06E-06	AT5G57035		1.01E-06	u-box domain-containing protein kinase family protein; phosphorylation
Latitude	TRINITY_DN17249_c1_g1_i2	5.08E-06	AT2G06210	ELF8*	4.28E-43	involved in control of flowering time by elevating FLC expression to a level that creates the vernalization responsive, winter annual habit; flower development; early flowering
Latitude	TRINITY_DN17796_c0_g2_i1	6.83E-06	AT1G80230		2.73E-13	rubredoxin-like superfamily protein; mitochondrial ATP transport
Latitude	TRINITY_DN8492_c0_g2_i1	9.52E-06	AT3G04780		7.58E-26	unknown protein
Latitude	TRINITY_DN13277_c2_g1_i1	1.03E-05	AT5G56860	GNC; GATA	6.86E-15	member of GATA factor family of zinc transcription factors; modulate chlorophyll biosynthesis and glutamate synthase expression
Latitude	TRINITY_DN11956_c1_g1_i3	1.42E-05	AT5G13850	NACA3	4.33E-43	nascent polypeptide-associated complex subunit alpha-like protein 3; protein transport
Latitude	TRINITY_DN18060_c0_g2_i2	1.69E-05	AT3G01770	BET10	7.15E-04	bromodomain protein that functions as a negative regulator of sugar and ABA signaling
Latitude	TRINITY_DN15276_c2_g3_i1	2.34E-05	AT1G11720	SS3	6.54E-07	starch biosynthesis; mutants have excess starch in leaves
Latitude	TRINITY_DN15488_c0_g1_i3	2.41E-05	AT1G65840	PAO4	1.69E-10	involved in back-conversion polyamine degradation pathway; major isoform in root peroxisomes
Latitude	TRINITY_DN16547_c0_g1_i1	3.67E-05	AT1G25350	OVA9	4.57E-107	glutamine-tRNA ligase; plant ovule development

Latitude	TRINITY_DN9090_c0_g1_i2	8.41E-05	AT5G30495		2.69E-10	Fcf2 pre-rRNA processing protein
Latitude	TRINITY_DN17587_c0_g2_i1	0.0001205	AT1G07010	SLP1	7.98E-12	Calcineurin-like metallo-phosphoesterase superfamily protein
Latitude	TRINITY_DN12124_c0_g1_i1	0.0001557	AT3G10230	LYC	8.04E-28	carotene biosynthesis
Latitude	TRINITY_DN15279_c0_g2_i2	0.0001722	AT1G55690		3.41E-24	protein transport; phosphatidylinositol transfer family protein
Latitude	TRINITY_DN11421_c0_g1_i1	0.0001958	AT1G32990	PRPL11	4.63E-52	mutant has decreased effective quantum yield of photosystem II; pale green plants
Latitude	TRINITY_DN16757_c0_g4_i2	0.0002136	AT2G42590	GRF9	7.39E-36	binds calcium and displays induced structural changes
Latitude	TRINITY_DN17122_c0_g1_i2	0.0002274	AT5G13640	PDAT	7.73E-47	glycerol metabolic process phospholipid diacylglycerol acyltransferase; seed viability
Latitude	TRINITY_DN18146_c0_g1_i2	0.0003063	AT5G35840	PHYC*	7.69E-42	one of a family of photoreceptors that modulate plant growth and development; detection of light; mutants flower late under LD compared with wildtype; flowering time
Latitude	TRINITY_DN12450_c0_g4_i1	0.0003928	AT4G33240	FAB1A	1.20E-16	pollen development; phosphatidylinositol phosphorylation; vacuole organization
Latitude	TRINITY_DN9713_c0_g2_i1	0.0004943	AT3G07590	SMD1A	1.19E-42	posttranscriptional gene silencing by RNA
Latitude	TRINITY_DN18131_c1_g1_i3	0.0004943	AT4G23230	CRK15	1.83E-12	cysteine-rich receptor like protein kinase; defense in response to bacterium; protein phosphorylation
Latitude	TRINITY_DN15831_c1_g3_i1	0.0005092	AT2G38360	PRA1.B4	2.13E-08	prenylated RAB acceptor; vesicle mediated transport
Latitude	TRINITY_DN17940_c1_g1_i4	0.0005589	AT2G41900	OXS2*	1.02E-24	regulation of transcription; stress tolerance
Latitude	TRINITY_DN18080_c0_g1_i8	0.0005661	AT3G14270	FAB1B	9.67E-76	pollen development, stomatal closure, vacuole organization; phosphatidylinositole-3 phosphate kinase
Latitude	TRINITY_DN15047_c0_g1_i1	0.0005805	AT1G54100	ALDH7B4*	7.58E-59	aldehyde dehydrogenase; response to desiccation, response to salt stress, response

						to abscisic acid
Latitude	TRINITY_DN13507_c0_g1_i1	0.000593	AT3G57300	INO8O*	4.53E-45	genome stability maintenance; flowering time; DNA repair; homologous recombination
Latitude	TRINITY_DN16118_c0_g1_i1	0.000593	AT3G14980	ROS4	1.22E-09	DNA demethylation regulation; histone acetyletransferase
Latitude	TRINITY_DN18101_c0_g2_i1	0.0006606	AT1G17210	ILP1	6.18E-20	IAP-like protein
Latitude	TRINITY_DN16628_c0_g2_i1	0.0007113	AT5G59700		1.92E-29	protein kinase superfamily protein
Longitude	TRINITY_DN10881_c0_g1_i1	2.64E-06	AT1G12840	DET3	6.39E-34	proton transmembrane transport; exocytosis; dark-grown mutant seedlings have light-grown morphology, develop true leaves and no chloroplasts, when grown in light plants are short and have reduced apical dominance
Longitude	TRINITY_DN18288_c2_g1_i2	4.52E-05	AT1G03080	NET1D	5.12E-34	kinase interacting family protein
Longitude	TRINITY_DN15103_c0_g1_i1	0.0001712	AT1G55250	HUB2	9.92E-16	encodes one of two orthologous ubiquitin ligases involved in monoubiquitination of histone; seed dormancy process; immune response; vegetative to reproductive transition of meristem; mutants have reduced seed dormancy phenotype (Liu et al. 2007), mutants had thinner inflorescence stems (Fleury et al. 2007); mutants have accelerated flowering compared with wildtype under LD and SD (Cao et al. 2008)
Longitude	TRINITY_DN14689_c0_g1_i1	0.0002465	AT1G78390	NCED9*	4.82E-13	biosynthesis of abscisic acid; expression increases during the first 6 hours of imbibing; seed dormancy + germination
Longitude	TRINITY_DN13579_c0_g2_i1	0.0002902	AT2G14750	APK	1.23E-25	essential for pollen viability; male gamete generation; sulfation of secondary metabolites including glucosinolates

Longitude	TRINITY_DN13821_c2_g1_i3	0.000352	AT4G25030	ATNHR2B*	8.01E-15	mutants susceptible to *P.syringae* and produce less callose upon infection; defense in response to bacterium
Longitude	TRINITY_DN15639_c2_g1_i2	0.0004679	AT4G27430	CIP7	2.39E-10	regulator of light-regulated genes; requires light for high expression; response to light stimulus
Longitude	TRINITY_DN18146_c0_g1_i2	0.0027092	AT5G35840	PHYC*	7.69E-42	photoreceptors that modulates plant growth and development; detection of visible light, far-red light phototransduction; mutants have late flowering under LD compared with wildtype mutants are elongated compared with wt under red light
Longitude	TRINITY_DN15781_c0_g3_i6	0.00303	AT3G50210		5.30E-22	iron dependent oxygenase superfamily protein; aging; cellular response to aging
Longitude	TRINITY_DN18090_c2_g2_i1	0.00303	AT5G06600	UBP12*	1.67E-174	ubiquitin-specific protease; jasmonic acid mediated signaling pathway; protein deubiquination; stress response
Longitude	TRINITY_DN18270_c1_g2_i2	0.0034	AT2G36590		9.58E-15	encodes a proline transported with affinity for gly betaine, proleine and GABA
Longitude	TRINITY_DN18479_c5_g2_i2	0.0034	AT3G62660	GATL7	1.39E-13	encodes protein with galacturonosyltransferase activity
Longitude	TRINITY_DN7556_c0_g1_i1	0.0045231	AT1G11260	STP1	6.72E-19	hexose cotransporter; monosaccharide transmembrane transport
Longitude	TRINITY_DN18506_c1_g1_i2	0.0050534	AT2G27150	AAO3	1.21E-08	involved in final step of abscisic acid biosynthesis; mutants are ABA deficient (Seo et al. 2004); mutants germinate after far red/red light treatment, whereas wildtype germination was suppressed (Seo et al. 2006); mutants have increased susceptibility to *P.irregulare* (Adie et al. 2007)
Longitude	TRINITY_DN14609_c0_g2_i2	0.0057709	AT3G19720		1.25E-48	mutants have defects in chloroplasts (Gao et

						al. 2003); endocytosis, mitochondrial division; circadian rhythm
Longitude	TRINITY_DN16031_c0_g3_i2	0.0064568	AT5G51230	EMF2	9.68E-34	involved in negative regulation of reproductive development; flower development; mutants have infertile flowers or incomplete floral development, short hypocotyls
Longitude	TRINITY_DN11410_c0_g2_i2	0.0069375	AT5G67370	CGLD27	1.58E-09	response to iron ion starvation; mutants shows stronger growth defect in low Fe (Urzica et al. 2012)
Longitude	TRINITY_DN17156_c0_g5_i2	0.0071667	AT3G03380	DEG7	6.29E-70	protease; photoinhibition; PSII repair
Longitude	TRINITY_DN12728_c1_g1_i3	0.0082029	AT4G21670	CPL1*	5.11E-52	transcriptional repressor; mutants exhibit hyperresponsiveness to ABA, cold and NaCl; response to salt stress; response to wounding; abscisic acid-activated signaling pathway
Longitude	TRINITY_DN17963_c1_g1_i4	0.0082029	AT1G11060	WAPL1	8.28E-07	involved in prophase removal of cohesion during meiosis; mutants exhibit reduced fertility and defects in embryo development; embryo development ending in seed dormancy;
Longitude	TRINITY_DN9859_c0_g1_i1	0.0082029	AT4G04950	GRXS17	9.40E-21	involved in ROS accumulation, auxin signaling and temperature-dependent postembryonic growth in plants; cellular response to DNA damage stimulus; response to heat; drought tolerance
Longitude	TRINITY_DN14567_c0_g1_i1	0.0082029	AT4G16580		7.82E-22	protein phosphatase family protein
Longitude	TRINITY_DN10679_c0_g1_i1	0.0121965	AT2G27480		3.83E-10	calcium-binding EF hand family protein
Longitude	TRINITY_DN18152_c0_g1_i2	0.0127081	AT1G10950	TMN1	7.76E-127	transmembrane protein; involved in cell adhesion and phagocytosis

Longitude	TRINITY_DN16286_c1_g1_i4	0.0131868	AT5G03190	CPuORF47	6.48E-13	peptide upstream protein;
Longitude	TRINITY_DN18181_c1_g2_i1	0.0131868	AT1G70160		1.47E-44	zinc finger MYND domain protein
Longitude	TRINITY_DN10307_c0_g1_i2	0.014003	AT5G67500	VDAC2	8.67E-08	encodes voltage dependent ion channel; involved in metabolite exchange between the organelle and cytosol; regulation of growth; response to bacterium;
Longitude	TRINITY_DN13844_c0_g4_i1	0.0141452	AT1G13190		1.01E-22	RNA-binding family protein
Longitude	TRINITY_DN13194_c0_g3_i1	0.0141452	AT4G35350	XCP1	1.07E-19	tracheary element vacuolar protein
Longitude	TRINITY_DN15657_c1_g1_i3	0.0141452	AT5G50250	CP31B	4.27E-16	RNA binding protein; RNA modification; innate immune response
Longitude	TRINITY_DN7019_c0_g1_i1	0.0141593	AT1G11390		8.76E-31	protein kinase superfamily protein
Longitude	TRINITY_DN16434_c0_g3_i1	0.0141593	AT2G01570	RGA1	6.39E-07	member of VHIID/DELLA regulatory family; transcriptional regulator; DELLAs repress cell proliferation and expansion that drives plant growth; involved in fruit and flowering development; represses GA-induced vegetative growth and floral initiation
Longitude	TRINITY_DN12767_c0_g1_i1	0.0141593	AT1G13170	ORP1D	2.48E-108	oxysterol binding protein
Longitude	TRINITY_DN18116_c1_g3_i2	0.0141593	AT1G68370	ARG1	5.97E-102	protein with homology to coiled coil found in cytoskeleton-interacting proteins; positive gravitropism
Longitude	TRINITY_DN13541_c0_g6_i1	0.0141593	AT1G23080	PIN7	1.10E-37	involved during embryogenesis; involved in pattern specification during root development; auxins signaling pathway;
Longitude	TRINITY_DN16349_c2_g1_i1	0.0141593	AT1G61150		1.28E-40	LisH and RanBPM domains containing protein
Longitude	TRINITY_DN10393_c0_g1_i1	0.0157174	AT1G58080	ATP-PRT1	9.21E-33	ATP phosphoribosyl transferase; histidine biosynthesis

Longitude	TRINITY_DN15674_c0_g1_i1	0.0157174	AT1G76630	SKI3	7.52E-11	involved in exosome mediated RNA decay; postranscriptional gene silencing; wax biosynthesis
Longitude	TRINITY_DN14913_c0_g1_i2	0.01588	AT2G24590	RSZ22A	6.89E-10	protein splicing factor; RNA splicing, regulation of mRNA splicing
Flowering time	TRINITY_DN15479_c1_g1_i1	0.00039	AT1G19110		3.10E-06	inter-alpha-trypsin inhibitor heavy chain-like protein
Flowering time	TRINITY_DN16875_c3_g1_i2	0.0015566	AT4G34310		1.87E-48	alpha/beta hydrolases superfamily protein
Flowering time	TRINITY_DN18299_c0_g1_i3	0.0015566	AT5G13100		4.37E-24	gap junction beta-4 protein
Flowering time	TRINITY_DN16515_c0_g1_i1	0.0021572	AT1G77620		2.94E-04	p-loop containing nucleoside tri-phosphate hydrolases superfamily protein; DNA repair
Flowering time	TRINITY_DN18479_c5_g4_i1	0.0021572	AT1G70090	LGT8	1.07E-13	protein with putative galacturonosyltransferase activity
Flowering time	TRINITY_DN13443_c3_g3_i1	0.0035573	AT4G12610	ATRAP74	8.74E-67	transcription initiation factor
Flowering time	TRINITY_DN17249_c1_g1_i2	0.0042212	AT2G06210	ELF8*; VIP6	4.28E-43	involved in control of flowering time by elevating FLC expression to a level that creates the vernalization-responses, winter annual habit; early flowering
Flowering time	TRINITY_DN12923_c0_g1_i1	0.0058662	AT4G36530		2.49E-33	alpha beta hydrolases superfamily protein
Flowering time	TRINITY_DN17844_c1_g2_i1	0.0073148	AT2G32390	ATGLR3.5*	4.59E-21	encodes an ionotropic glutamate receptor ortholog; member of ligand-gated ion channel subunit family; stomatal movement
Flowering time	TRINITY_DN15849_c0_g1_i1	0.0174505	AT5G23430	KTN80.4	6.45E-48	one of four katanin p80 subunits; severing of microtubules
Flowering time	TRINITY_DN12216_c0_g1_i2	0.0213467	AT1G44170	ALDH3H1	4.15E-29	aldehyde dehydrogenase; induced by ANA and dehydration that can oxidize saturated aliphatic aldehydes

Flowering time	TRINITY_DN17871_c0_g2_i1	0.0213467	AT1G32440	Pkp3	1.42E-72	seed oil biosynthesis; chloroplast pyruvate kinase
Flowering time	TRINITY_DN13576_c0_g2_i1	0.0272861	AT1G53050		1.21E-14	protein kinase superfamily protein
Flowering time	TRINITY_DN10263_c0_g1_i1	0.0298187	AT5G05410	DREB2A*	1.81E-12	encode transcription factor that binds to DRE/CRT cis elements responsive to drought and low-temperature stress; over expression resulted in drought stress tolerance but only slightly freezing tolerance; regulates expression of many water stress inducible genes
Flowering time	TRINITY_DN12652_c0_g1_i1	0.0298187	AT2G23470	RUS4	3.03E-30	root UVB sensitive protein
Flowering time	TRINITY_DN13277_c2_g1_i1	0.0298187	AT5G56860	GNC	6.86E-15	modulate chlorophyll biosynthesis and glutamate synthase expression; member of GATA factor family of zinc finger transcription factors; reduced chlorophyll levels; greening
Flowering time	TRINITY_DN14423_c1_g2_i2	0.0298187	AT5G65630	GTE7	3.98E-04	involved in resistance to agrobacterium-mediated root transformation; encodes a bromodomain-containing protein
Flowering time	TRINITY_DN16330_c1_g1_i1	0.0298187	AT2G40970	MYBC1	1.03E-16	homeodomain-like superfamily protein; freezing tolerance
Flowering time	TRINITY_DN17256_c0_g1_i1	0.0298187	AT4G25730	TRM7B	1.02E-21	methyltransferase family protein
Flowering time	TRINITY_DN16133_c0_g2_i6	0.0306815	AT4G21270	ATK1	2.80E-34	loss of function mutations have reduced fertility and are defective in spindle formation in male meiosis; kinsesin-like motor protein heavy chain
Flowering time	TRINITY_DN12710_c0_g1_i1	0.0319263	AT5G20080		4.48E-16	FAD/NAPP binding oxidoreductase

Flowering time	TRINITY_DN7156_c0_g1_i1	0.0324759	AT5G59140		2.16E-19	BTB/POZ domain-containing protein
Flowering time	TRINITY_DN15582_c2_g1_i1	0.0324759	AT5G51830	FRK1	5.26E-19	fructokinase; important for seed oil accumulation and vascular development
Flowering time	TRINITY_DN14029_c2_g5_i1	0.032564	AT1G32330	HSFA1D*	4.42E-17	member of heat stress transcription factor family; response to heat
Flowering time	TRINITY_DN12683_c1_g1_i1	0.0340632	AT5G24930	COL4*	2.88E-20	flowering repressor in long day and short days and acts on the expression of FT and FT-like genes as well as on SOC1
Flowering time	TRINITY_DN16514_c0_g1_i2	0.0367164	AT5G13190	GILP	1.23E-15	negative regulation of hypersensitive cell death; plasma membrane protein
Flowering time	TRINITY_DN17695_c1_g4_i3	0.0395247	AT5G63410		1.28E-04	leucine-rich repeat protein kinase family
Flowering time	TRINITY_DN17625_c4_g2_i1	0.0395247	AT5G36250	PP2C74	8.00E-24	protein phosphatase
Flowering time	TRINITY_DN11608_c0_g1_i1	0.0426213	AT2G02860	SUT2*	1.57E-38	gene expression is induced by wounding; response to wounding; sucrose transport
Flowering time	TRINITY_DN15172_c0_g1_i1	0.0426213	AT5G28530	FRS10*	3.09E-15	FAR1-related sequence; light control
Flowering time	TRINITY_DN16708_c0_g1_i2	0.0426213	AT1G10410		1.47E-10	CW14 protein
Flowering time	TRINITY_DN16822_c1_g5_i2	0.0426213	AT4G16280.1	FCA	5.76E-14	flowering control locus A; involved in the transition of the vegetative meristem to reproduction development; late flowering
Flowering time	TRINITY_DN16018_c0_g1_i1	0.0444423	AT4G32285.1	CAP1	1.22E-11	clathrin mediated endocytosis
Flowering time	TRINITY_DN15313_c0_g2_i1	0.0446878	AT2G42990.1		3.36E-04	lipid catabolic process; GDSL-motif esterase/acetyltrasnferase/lipase
Flowering time	TRINITY_DN16320_c0_g1_i1	0.0446878	AT1G68570.2	NPF3.1	4.35E-08	membrane localized GA transported that is expressed in root endodermis

Flowering time	TRINITY_DN12230_c0_g1_i1	0.0447871	AT5G09260.1	VPS20.2	1.49E-43	vacuolar protein sorting-associated protein; protein transport
Flowering time	TRINITY_DN17034_c0_g1_i1	0.0447871	AT3G01120.1	MTO1	1.04E-26	methionine biosynthesis
Flowering time	TRINITY_DN17660_c0_g3_i1	0.0447871	AT5G13820.2	TBP1	4.71E-17	encodes a protein that binds to plant telomeric DNA repeats; telomere maintenance

Appendix 4

A. Supplementary figures

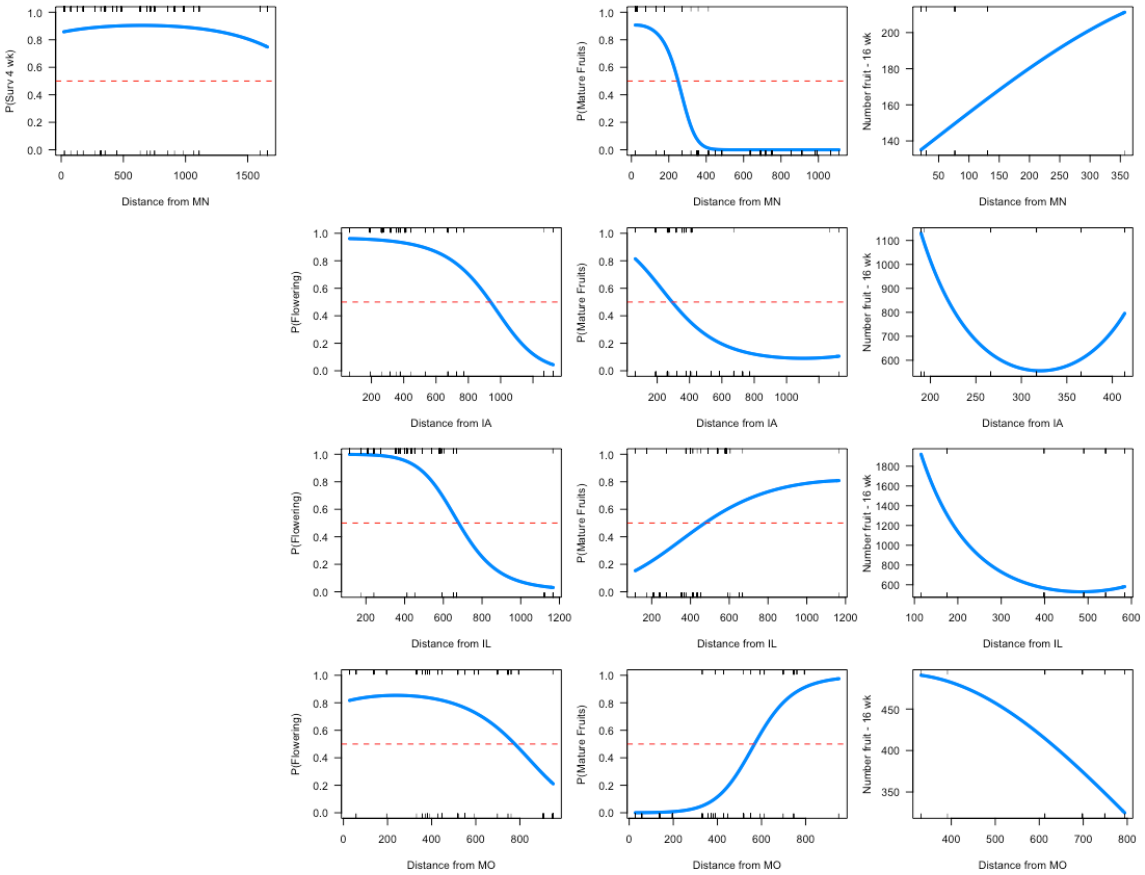

Figure A1. The quadratic relationship between fitness and geographic distance of source population from each of the four transplant sites. Generalized linear mixed models were run on individual data and included block (random), geographic distance and geographic distance^2 from field site as predictors. The model outcomes were transformed back to the response scale. Each row is a different field site (MN, IA, IL, and MO), and each column is a different fitness component (P(Surv 4 wk), P(Flowering), P(Mature Fruit), Number of Fruit at 16 weeks). The blue line represents the expected value, and red dotted line indicates where the probability is 0.5. There are no confidence intervals due to the inclusion of block as a random effect. Blank plots indicate those fitness components whereby the model would not run.

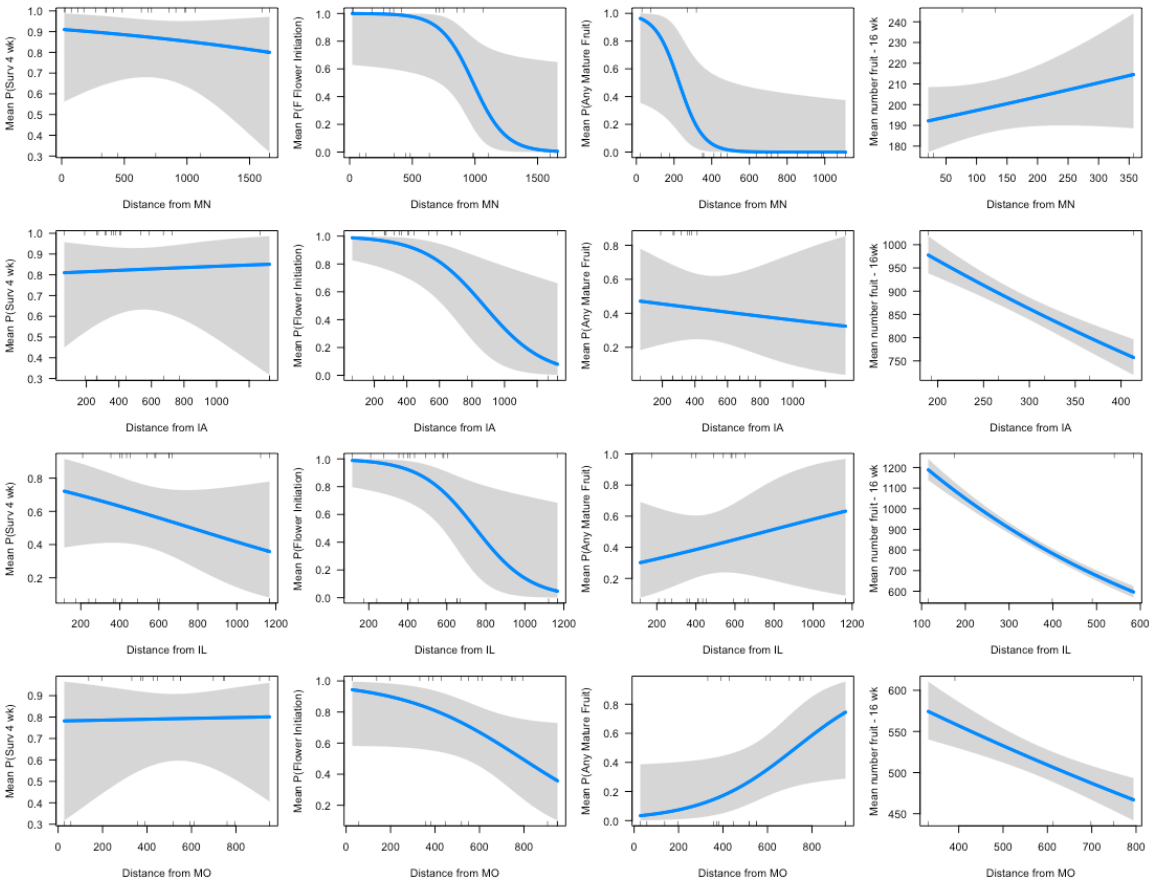

Figure A2. The relationship between population mean fitness and geographic distance of source population from each of the four transplant sites. Generalized linear models were run on population means with geographic distance from field site as a predictor. The model outcomes were transformed back to the response scale. Each row is a different field site (MN, IA, IL, and MO), and each column is a different fitness component (P(Surv 4 wk), P(Flowering), P(Mature Fruit), Number of Fruit at 16 weeks). The blue line represents the expected value, the grey band indicates the confidence intervals around the expected value.

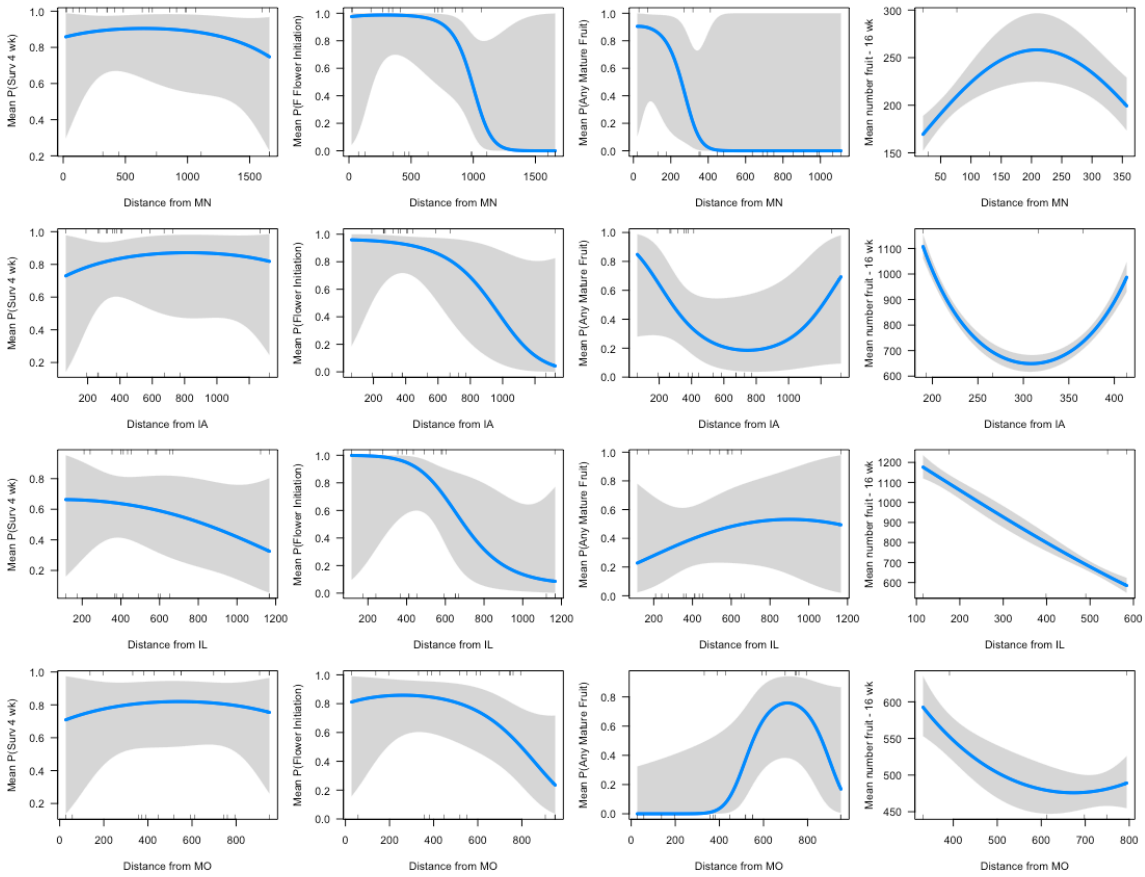

Figure A3. The quadratic relationship between population mean fitness and geographic distance of source population from each of the four transplant sites. Generalized linear models were run on population means with geographic distance and geographic distance^2 as predictors. The model outcomes were transformed back to the response scale Each row is a different field site (MN, IA, IL, and MO), and each column is a different fitness component (P(Surv 4 wk), P(Flowering), P(Mature Fruits), Number of Fruit at 16 weeks). The blue line represents the expected value, the grey band indicates the confidence intervals around the expected value.

B. Supplementary tables

Table A1. The effect of an individual's geographic distance from field site on conditional fitness components. Geographic distance (fixed), geographic distance^2 (fixed), and block (random) were included as predictors. Fixed effects were evaluated with type II Wald chi-square test, while random effects were evaluated with LRT comparing models with and without block (LRT df = 1). Residual deviance and df were obtained from the model without block. Models which would not run have blank rows. The results of this model correspond to Figure A1.

Fitness components	Factor	df	χ^2 or deviance	P-value
MN				
P(Survival)	Geographic distance	1	53.1	**<0.0001**
	Geographic distance^2	1	1.9	0.17
	Block		0	1
	Residual deviance	646	490.9	
P(Flowering)	Geographic distance			
	Geographic distance^2			
	Block			
	Residual deviance			
P(Mature fruit)	Geographic distance	1	0.48	0.49
	Geographic distance^2	1	511.2	**<0.0001**
	Block		0	1
	Residual deviance	407	148.9	
Number of fruits – 16 wks	Geographic distance	1	61.9	**<0.0001**
	Geographic distance^2	1	1.7	0.19
	Block		6637	**<0.0001**
	Residual deviance	53	9254	
IA				
P(Survival)	Geographic distance			
	Geographic distance^2			
	Block			
	Residual deviance			
P(Flowering)	Geographic distance	1	0.9	0.35
	Geographic distance^2	1	9.3	**0.002**
	Block		0	1
	Residual deviance	531	299.5	
P(Mature fruit)	Geographic distance	1	144.4	**<0.0001**
	Geographic distance^2	1	7.8	**0.005**
	Block		0	**1**
	Residual deviance	411	499.2	
Number of fruits – 16 wks	Geographic distance	1	2616	**<0.0001**
	Geographic distance^2	1	2270	**<0.0001**
	Block		12594	**<0.0001**
	Residual deviance	69	30049	
IL				
P(Survival)	Geographic distance			
	Geographic distance^2			
	Block			
	Residual deviance			
P(Flowering)	Geographic distance	1	138.3	**<0.0001**

	Geographic distance^2	1	9.9	**0,001**
	Block		0	1
	Residual deviance	368	212.7	
P(Mature fruit)	Geographic distance	1	52.4	**<0.0001**
	Geographic distance^2	1	2.7	0.10
	Block		0	1
	Residual deviance	290	373.9	
Number of fruits – 16 wks	Geographic distance	1	1661	**<0.0001**
	Geographic distance^2	1	939.8	**<0.0001**
	Block		8580	**<0.0001**
	Residual deviance	63	22841	
MO				
P(Survival)	Geographic distance			
	Geographic distance^2			
	Block			
	Residual deviance			
P(Flowering)	Geographic distance	1	44.9	**<0.0001**
	Geographic distance^2	1	41.1	**<0.0001**
	Block		0	1
	Residual deviance	497	527.0	
P(Mature fruit)	Geographic distance	1	236.6	**<0.0001**
	Geographic distance^2	1	32.7	**<0.0001**
	Block		0	1
	Residual deviance	330	243.4	
Number of fruits – 16 wks	Geographic distance	1	4.1	**0.04**
	Geographic distance^2	1	16.2	**<0.0001**
	Block		6418.1	**<0.0001**
	Residual deviance	47	9595	

Table A2. The effect of geographic distance from field site on conditional fitness components, conducted on population means. Each fitness component was averaged for each population prior to running the analyses. Geographic distance was included as a predictor and was evaluated using a likelihood ratio test comparing nested models. The results of this model correspond to Figure A2.

Fitness components	Factor	df	Deviance	P-value
MN				
P(Survival)	Geographic distance	1	0.3	0.61
	Residual deviance	24	5.5	
P(Flowering)	Geographic distance	1	19.7	**<0.0001**
	Residual deviance	24	2.4	
P(Mature fruit)	Geographic distance	1	16.1	**<0.0001**
	Residual deviance	20	1.4	
Number of fruits – 16 wks	Geographic distance	1	1.7	0.19
	Residual deviance	3	21.2	
IA				
P(Survival)	Geographic distance	1	0.02	0.87
	Residual deviance	24	2.9	
P(Flowering)	Geographic distance	1	12.9	**<0.0001**
	Residual deviance	24	3.3	
P(Mature fruit)	Geographic distance	1	0.1	0.72
	Residual deviance	22	22.8	
Number of fruits – 16 wks	Geographic distance	1	47.1	**<0.0001**
	Residual deviance	4	395.2	
IL				
P(Survival)	Geographic distance	1	1.2	0.28
	Residual deviance	24	2.6	
P(Flowering)	Geographic distance	1	12.9	**<0.0001**
	Residual deviance	24	8.0	
P(Mature fruit)	Geographic distance	1	0.4	0.51
	Residual deviance	21	20.2	
Number of fruits – 16 wks	Geographic distance	1	361.3	**<0.0001**
	Residual deviance	4	674.4	
MO				
P(Survival)	Geographic distance	1	0.005	0.95
	Residual deviance	24	3.9	
P(Flowering)	Geographic distance	1	4.2	**<0.0001**
	Residual deviance	24	13.8	
P(Mature fruit)	Geographic distance	1	5.2	**0.02**
	Residual deviance	22	15.8	
Number of fruits – 16 wks	Geographic distance	1	19.3	**<0.0001**
	Residual deviance	4	80.7	

Table A3. The effect of geographic distance from field site on conditional fitness components, conducted on population means. Each fitness component was averaged for each population prior to running the analyses. Geographic distance and geographic distance^2 were included as predictors. All predictors were evaluated using a likelihood ratio test comparing nested models. The results of this model correspond to Figure A3.

Fitness components	Factor	df	Deviance	P-value
MN				
P(Survival)	Geographic distance	1	0.1	0.72
	Geographic distance^2	1	0.3	0.61
	Residual deviance	23	5.2	
P(Flowering)	Geographic distance	1	0.2	0.70
	Geographic distance^2	1	0.8	0.38
	Residual deviance	23	1.6	
P(Any fruit)	Geographic distance	1	0.0	0.94
	Geographic distance^2	1	0.3	0.56
	Residual deviance	19	1.0	
Number of fruits – 16 wks	Geographic distance	1	15.6	**<0.0001**
	Geographic distance^2	1	14.2	**0.0001**
	Residual deviance	2	6.9	
IA				
P(Survival)	Geographic distance	1	0.2	0.68
	Geographic distance^2	1	0.2	0.70
	Residual deviance	23	2.8	
P(Flowering)	Geographic distance	1	0.0	0.97
	Geographic distance^2	1	0.3	0.57
	Residual deviance	23	3.0	
P(Any fruit)	Geographic distance	1	3.3	0.07
	Geographic distance^2	1	3.2	0.07
	Residual deviance	21	19.6	
Number of fruits – 16 wks	Geographic distance	1	212.4	**<0.0001**
	Geographic distance^2	1	194.2	**<0.0001**
	Residual deviance	3	201.1	
IL				
P(Survival)	Geographic distance	1	0.0	0.97
	Geographic distance^2	1	0.1	0.77
	Residual deviance	23	2.6	
P(Flowering)	Geographic distance	1	1.1	0.25
	Geographic distance^2	1	0.4	0.54
	Residual deviance	23	7.5	
P(Any fruit)	Geographic distance	1	0.4	0.55
	Geographic distance^2	1	0.2	0.67
	Residual deviance	20	20.1	
Number of fruits – 16 wks	Geographic distance	1	3.5	0.06
	Geographic distance^2	1	0.8	0.37
	Residual deviance	3	673.7	
MO				
P(Survival)	Geographic distance	1	0.2	0.66
	Geographic distance^2	1	0.2	0.66

	Residual deviance	23	3.7	
P(Flowering)	Geographic distance	1	0.2	0.65
	Geographic distance^2	1	0.9	0.34
	Residual deviance	23	12.8	
P(Any mature fruit)	Geographic distance	1	8.5	**0.003**
	Geographic distance^2	1	6.7	**0.009**
	Residual deviance	21	8.9	
Number of fruits – 16 wks	Geographic distance	1	5.1	**0.02**
	Geographic distance^2	1	3.5	**0.06**
	Residual deviance	3	77.2	

Table A4. The effect of latitudinal distance from field site on conditional fitness components, conducted on population means. Each fitness component was averaged for each population prior to running the analyses. Latitudinal distance was included as a predictor and was evaluated using a likelihood ratio test comparing nested models.

Fitness components	Factor	df	Deviance	P-value
MN				
P(Survival)	Latitude distance	1	0.2	0.70
	Residual deviance	24	5.6	
P(Flowering)	Latitude distance	1	19.6	**<0.0001**
	Residual deviance	24	2.5	
P(Any mature fruit)	Latitude distance		17.0	**<0.0001**
	Residual deviance	20	2.3	
Number of fruits – 16 wks	Latitude distance	1	0.1	0.80
	Residual deviance	3	22.8	
IA				
P(Survival)	Latitude distance	1	0.1	0.82
	Residual deviance	24	2.9	
P(Flowering)	Latitude distance	1	11.5	**0.0007**
	Residual deviance	24	4.8	
P(Mature fruit)	Latitude distance	1	5.0	0.02
	Residual deviance	22	17.9	
Number of fruits – 16 wks	Latitude distance	1	68.6	**<0.0001**
	Residual deviance	4	373.7	
IL				
P(Survival)	Latitude distance	1	0.5	0.47
	Residual deviance	24	3.2	
P(Flowering)	Latitude distance	1	16.8	**<0.0001**
	Residual deviance	24	4.4	
P(Mature fruit)	Latitude distance	1	9.9	0.002
	Residual deviance	21	10.5	
Number of fruits – 16 wks	Latitude distance	1	354.3	**<0.0001**
	Residual deviance	4	681.5	
MO				
P(Survival)	Latitude distance	1	0.1	0.69
	Residual deviance	24	3.7	
P(Flowering)	Latitude distance	1	14.1	**0.0001**
	Residual deviance	24	3.8	
P(Mature fruit)	Latitude distance	1	20.5	**<0.0001**
	Residual deviance	22	0.7	
Number of fruits – 16 wks	Latitude distance	1	18.5	**<0.0001**
	Residual deviance	4	81.6	

Table A5. The effect of spatial scale on conditional fitness components, conducted on population means. Each fitness component was averaged for each population prior to running the analyses. 'Zone' refers to the USDA hardiness zone for the location of each population, 'region' refers to the spatial grouping within a given hardiness zone, 'population' refers to the individual populations found within a given regional grouping. All predictors were evaluated using a likelihood ratio test comparing nested models.

Fitness components	Factor	df	Deviance	P-value
MN				
P(Survival)	Climate zone	3	0.45	0.93
	Region(Climate zone)	6	1.74	0.94
	Population(Region(Climate zone))	111	2.33	0.99
P(Flowering)	Climate zone	3	8.95	**0.03**
	Region(Climate zone)	6	0.16	0.99
	Population(Region(Climate zone))	11	0.59	1.00
P(Mature fruit)	Climate zone	3	14.5	**0.002**
	Region(Climate zone)	6	1.72	0.94
	Population(Region(Climate zone))	11	0.63	1.00
IA				
P(Survival)	Climate zone	3	0.33	0.95
	Region(Climate zone)	6	1.15	0.98
	Population(Region(Climate zone))	11	1.34	0.99
P(Flowering)	Climate zone	3	1.31	0.72
	Region(Climate zone)	6	0.76	0.99
	Population(Region(Climate zone))	11	1.67	0.99
P(Mature fruit)	Climate zone	3	18.7	**0.0003**
	Region(Climate zone)	6	1.55	0.96
	Population(Region(Climate zone))	11	0.87	0.99
IL				
P(Survival)	Climate zone	3	0.25	0.97
	Region(Climate zone)	6	1.41	0.97
	Population(Region(Climate zone))	11	0.90	1.00
P(Flowering)	Climate zone	3	9.46	**0.02**
	Region(Climate zone)	6	0.28	0.99
	Population(Region(Climate zone))	11	0.92	0.99
P(Mature fruit)	Climate zone	3	15.9	**0.001**
	Region(Climate zone)	6	2.15	0.90
	Population(Region(Climate zone))	11	2.19	0.99
MO				
P(Survival)	Climate zone	3	0.69	0.88
	Region(Climate zone)	6	1.27	0.97
	Population(Region(Climate zone))	11	1.63	0.99
P(Flowering)	Climate zone	3	6.36	0.09
	Region(Climate zone)	6	0.42	0.99
	Population(Region(Climate zone))	11	1.29	0.99
P(Mature fruit)	Climate zone	3	19.1	**0.0002**

Region(Climate zone)	6	0.10	0.99
Population(Region(Climate zone))	11	0.41	0.99